JOURNAL OF COMPUTER AND SYSTEM SCIENCES **53**, 469–486 (1996)
ARTICLE NO. 0081

A Threshold for Unsatisfiability

Andreas Goerdt

Fakultät für Informatik, Theoretische Informatik/TU Chemnitz-Zwickau, D-09111 Chemnitz, Germany

Received August 12, 1991; revised May 20, 1996

A propositional formula is in 2-CNF (2-conjunctive normalform) iff it is the conjunction of clauses each of which has exactly two literals. We show: If $C = 1 + \varepsilon$, where $\varepsilon > 0$ is fixed and $q(n) \geqslant C \cdot n$, then almost all formulas in 2-CNF with $q(n)$ different clauses, where n is the number of variables, are unsatisfiable. If $C = 1 - \varepsilon$ and $q(n) \leqslant C \cdot n$, then almost all formulas with $q(n)$ clauses are satisfiable. By "almost all" we mean that the probability of the set of unsatisfiable or satisfiable formulas among all formulas with $q(n)$ clauses approaches 1 as $n \to \infty$. So $C = 1$ gives us a threshold separating satisfiability and unsatisfiability of formulas in 2-CNF in a probabilistic, asymptotic sense. To prove our result we translate the satisfiability problem for formulas in 2-CNF into a graph theoretical question. Then we apply techniques from the theory of random graphs.

INTRODUCTION

The motivation of this paper comes from the area of probabilistic analysis of satisfiability algorithms. The worst-case complexity $f(n)$ of an algorithm A is usually defined like $f(n) = \text{Max}\{\text{runtime of } A \text{ with input } I \mid I \text{ is presented by } n \text{ bits}\}$. Now it seems natural to define the average-case complexity g of A as

$$g(n) = \frac{\sum \text{runtime of } A \text{ with input } I}{|\{I \mid I \text{ is represented by } n \text{ bits}\}|},$$

where the sum goes over all inputs I of A of length n. However, this is usually not done. One reason for this is that among all inputs which can be presented by n bits there are often many trivial inputs which can be solved in a few steps by A. This does not affect the worst-case complexity but renders the above average complexity meaningless. Therefore special natural looking families of input spaces (F_n) are designed where each F_n often is a subset of all inputs of length n or some polynomial in n. Then the average complexity is computed with respect to such families of input spaces endowed with a probability distribution (usually the uniform one).

The main target of probabilistic analysis of satisfiability algorithms has been the Davis–Putnam procedure, a simple backtracking algorithm. In [2, 9] this procedure is analyzed when the family of input spaces is given by: $\text{Form}_n(q, k) =$ the set of all propositional formulas over n variables with $q = q(n)$ nontautological clauses having exactly k literals each. Multiple occurrences of the same clause are allowed and each formula from $\text{Form}_n(q, k)$ is equally likely. The purpose of this paper is to contribute to an analysis of these input spaces.

Naturally the main target of investigation is $\text{Form}_n(q, 3)$, because $k = 3$ is the smallest k such that no polynomial time satisfiability algorithm is known. For $k = 2$ we can test satisfiability in linear time [1]. If $q = q(n)$ approaches $C \cdot n$ for a constant C the following results are known about $\text{Form}_n(q, 3)$: The average size of the whole backtracking tree of a simplified Davis–Putnam procedure is exponential in n [2, 9]. (There are different families of input spaces where this size is polynomial [10].) For almost all *unsatisfiable* formulas from $\text{Form}_n(q, 3)$ the length of the shortest resolution proof is exponential, i.e, $\geqslant (1 + \varepsilon)^n$ for a fixed $\varepsilon > 0$ [5]. On the other hand, the following positive results are known: The Davis–Putnam procedure using only the pure literal heuristics finds a satisfying truth value assignment in polynomial time for almost all formulas of $\text{Form}_n(q, 3)$ if $C < 1$ [8]. This procedure with another heuristics finds a satisfying assignment in polynomial time for a portion of $\text{Form}_n(q, 3)$ whose probability is $\geqslant \varepsilon$ for a fixed $\varepsilon > 0$ if $C < 2.99$ [4]. Note that the size of C corresponds to the "degree of unsatisfiability."

Let X: $\text{Form}_n(q, 3) \to \mathcal{N}$ be the random variable assigning to each formula F the number of truth value assignments of the n underlying variables, which make F true. If $C > (-\ln 2)/(\ln 7/8)$ (≈ 5.19) we have for the expectation of X that $EX \to 0$, hence almost no formula is satisfiable [6, 7, 9]. If $C < (-\ln 2)/(\ln 7/8)$ then $EX \to \infty$. (See [6, 9] for additional experiments concerning EX.) But $EX \to \infty$ does not mean that almost all formulas are satisfiable. Many satisfying assignments can be concentrated in a small portion of all formulas. The computation of the variance of X and experiments in [3] show that this must be the case. The experiments suggest that for $C < 4$ the formulas tend to be satisfiable and for $C > 4$ unsatisfiable. A proof of this is not known.

It is generally believed and can be confirmed by experiments, that satisfiability algorithms when run with samples from the probability spaces $\text{Form}_n(q, 3)$ have a characteristic peak in their average running time when C is

0022-0000/96 $18.00

around 4 (recall $q(n)$ approaches $C \cdot n$). So formulas at the conjectured satisfiability threshold exhibit some kind of average case hardness. This makes them suitable as benchmarks for the average case behavior of satisfiability algorithms.

After the appearance of the first version of this paper [17] there has been considerable progress in the 3-CNF case. However a sharp threshold has not been proved [19–21].

The subject matter of this paper is the simpler space $\mathrm{Form}_n(q) = \mathrm{Form}_n(q, 2)$ with $q = q(n)$ approaching $C \cdot n$ for a constant C, modified such that *no* double occurrences of clauses are allowed. Looking at $\mathrm{Form}_n(q)$ *with* double occurrences it is known that $EX \to 0$ for $C > (-\ln 2)/(\ln 3/4)$ (≈ 2.49) and $EX \to \infty$ for $C < (-\ln 2)/(\ln 3/4)$, where X is the random variable as above. Experiments suggest a satisfiability threshold for $C = 1$ (cf. [12, Table 1, p. 228]).

To prove our threshold result we can build on the literature dealing with the peculiar properties of 2-CNF. An early paper in this area is [22]. We need the results of [1]. In [1] the question of unsatisfiability of a propositional formula in 2-CNF is reduced to the existence of certain cycles in certain graphs. Applying this technique to our random formulas allows us to reduce the question of unsatisfiability to the existence of a special kind of cycles in certain random graphs. Then we apply techniques from the theory of random graphs (the "first moment method" and the "second moment method" [14]) to obtain the threshold. As the spaces of random graphs considered here are different from the usual spaces of random graphs this application is not straightforward. The second moment method in its standard form is used to show the existence of subgraphs of *constant* size in almost all random graphs with a certain number of edges. Here we apply this method to show the existence of cycles of *increasing* size. We build on the methods introduced in [16].

A first version of this paper is [17], the same threshold has been obtained independently in [18].

In Section 1 we fix terminology and prove same basic results. In Section 2 we show that for $C < 1$ almost all formulas are satisfiable. Section 3 deals with the case $C > 1$.

1. BASICS

We consider propositional formulas over finite sets of propositional variables. Let $\mathrm{Var}_n = \{x_1, ..., x_n\}$ be a standard set of n propositional *variables*. $\mathrm{Lit}_n = \{x_1, ..., x_n, \bar{x}_1, ..., \bar{x}_n\}$ is the set of *literals* over Var_n. We let $\bar{\bar{x}} = x$. The set of *clauses* over Var_n is

$$\mathrm{Clause}_n = \{\{L, K\} \mid L, K \in \mathrm{Lit}_n, L \neq K, \bar{L} \neq K\}.$$

We write $L \vee K$ for $\{L, K\}$. Note that our definition neither allows unit clauses nor tautological clauses. We have not checked whether our results depend on this condition. (We do not think they do.) The reason for this condition is to reduce the technical effort in the subsequent calculations as much as possible. We have

$$|\mathrm{Clause}_n| = 4 \cdot \binom{n}{2} = 2n(n-1).$$

We let n always be the number of variables. $N = N(n) = 2n \cdot (n-1)$ is the number of clauses. We always assume that n is large. The set of *formulas* consisting of q clauses over n variables is given by

$$\mathrm{Form}_n(q) = \{\{C_1, ..., C_q\} \mid C_i \in \mathrm{Clause}_n, C_i \neq C_j \text{ for all } i \neq j\}.$$

For $\{C_1, ..., C_q\}$ we write $C_1 \wedge \cdots \wedge C_q$. We have

$$|\mathrm{Form}_n(q)| = \binom{N}{q}.$$

A mapping $\pi\colon \mathrm{Var}_n \to \{0, 1\}$ is a truth value assignment (0 for false and 1 for true). We extend π to Lit_n by letting $\pi(\bar{x}) = 1$ iff $\pi(x) = 0$. We say $\pi \models L \vee K$, π satisfies $L \vee K$, iff $\pi(L) = 1$ or $\pi(K) = 1$. Moreover $\pi \models C_1 \wedge \cdots \wedge C_q$, π *satisfies* $C_1 \wedge \cdots \wedge C_q$ iff $\pi \models C_i$ for all i.

We consider $\mathrm{Form}(q)$ as a probability space with the uniform distribution:

$$\Pr(F) = 1 \Big/ \binom{N}{q}$$

for $F \in \mathrm{Form}(q)$. $\mathrm{Sat}_n(q)$ is the set of satisfiable formulas from $\mathrm{Form}_n(q)$; $\mathrm{Unsat}_n(q)$ is the set of unsatisfiable formulas.

A *formula graph* over n variables is a *directed* graph $G = (V, E)$ with $V = \mathrm{Lit}_n$, $E \subseteq \{(L, K) \mid L, K \in V, L \neq K, L \neq \bar{K}\}$ and $(L, K) \in E \Leftrightarrow (\bar{K}, \bar{L}) \in E$. We call $L \to K$, $\bar{K} \to \bar{L}$ a pair of *complementary edges*. $FG_n(q)$ is the *set of formula graphs* over n variables with exactly q pairs of complementary edges. Let $F \in \mathrm{Form}_n(q)$. The *formula graph of* F, $FG(F) \in FG_n(q)$ is given by: $FG(F) = (V, E)$ with $V = \mathrm{Lit}_n$ and $\bar{L} \to K, \bar{K} \to L \in E$ iff $L \vee K$ is a clause of F. Note that $L \vee K$ is equivalent to $\bar{L} \to K \wedge \bar{K} \to L$. As the mapping $\{L, K\} \mapsto \{\bar{L} \to K, \bar{K} \to L\}$ is bijective we have

$$|FG_n(q)| = \binom{N}{q}.$$

$FG_n(q)$ becomes a probability space by defining

$$\Pr(G) = 1 \Big/ \binom{N}{q}.$$

$$\begin{array}{ccc} x & \overset{\leftarrow}{\rightarrow} & y \\ \downarrow\uparrow & & \downarrow\uparrow \\ \bar{y} & \overset{\leftarrow}{\rightarrow} & \bar{x} \end{array}$$

FIGURE 1

The *complete* formula graph over n variables is the graph containing all N possible pairs of complementary edges.

For example if $F=(x \vee y) \wedge (\bar{x} \vee y) \wedge (x \vee \bar{y}) \wedge (\bar{x} \vee \bar{y})$, then we get the formula graph from Fig. 1. Let $G=(V, E)$ be a formula graph. A *path* W in G is a subgraph $W=L_1 \to \cdots \to L_k$ where $L_i \neq L_j$ for all $i \neq j$ and $L_i \to L_{i+1} \in E$. If $k=0$, then $W=\varnothing$, the *empty* path, if $k=1$, then $W=L_1$. Length$(W)=0$ if $k=0$. Length$(W)=k-1$ if $k \geqslant 1$. A *cycle* π in G is a subgraph $\pi=L_1 \to L_2 \to \cdots \to L_k \to L_1$, where $k \geqslant 2$, $L_i \neq L_j$ for all $i \neq j$, and $L_i \to L_{i+1}$, $L_k \to L_1 \in E$. We define Length$(\pi)=k$. If $H=(U, F)$ with $U \subseteq \mathrm{Lit}_n$ is a directed graph, *not* necessarily a formula graph, we let $E(H)=F$, the set of edges of H, and $\mathrm{Lit}(H)=U$, the set of *literals* of H. Often we write $L \in H$ for $L \in \mathrm{Lit}(H)$. Finally let $\mathrm{Pairs}(H)=\{\{L \to K, \ \bar{K} \to \bar{L}\} \mid L \to K \in F\}$ be the set of pairs of edges necessary for a formula graph to have H as a subgraph. A pair $L, \bar{L} \in \mathrm{Lit}(H)$ is called a *contradictory pair* of H. The *complementary graph* of H, $\bar{H}=(\bar{U}, \bar{F})$ is given by $\bar{U}=\{\bar{L} \mid L \in U\}$ and $\bar{F}=\{\bar{L} \to \bar{K} \mid K \to L \in F\}$. A formula graph G has H as a subgraph iff it has $\bar{H}$ as a subgraph. For the notion "strong component" we refer to [1]. For two strong components C and D of a graph G we say D is a *successor* of C iff there is an edge in G leading from a literal of C to a literal of D. The ordering on the strong components of G is the transitive closure of this successor relation.

1.1. THEOREM. *Let F be a formula and $G=FG(F)$. The following three statements are equivalent.*

(a) *F is unsatisfiable.*

(b) *G contains a contradictory strong component, i.e., one with a contradictory pair.*

(c) *G contains a contradictory cycle.*

Proof. (a) $\Rightarrow$ (b) (cf. also [1]) From the definition of strong component we get, C is a contradictory strong component of G iff $C=\bar{C}$.

Now, assume that G contains no contradictary strong component. We construct a truth value assignment satisfying F as follows: First we assign truth values to all strong components of G, such that holds: If $C<D$ and C is assigned a 1, then D is assigned a 1. For this we use the following algorithm:

$$\begin{tikzcd} L \arrow[r] & U_1 \arrow[r] & K \arrow[dl] \\ & V_2 \arrow[ul] & \end{tikzcd}$$

FIGURE 2

$M := \{C \mid C$ is a maximal strong component of G w.r.t. the ordering of strong components$\}$
until $M=\varnothing$ do
  for all $C \in M$ do
    if C has not yet been assigned a truth value
      then
        assign C the value 1
        assign $\bar{C}$ the value 0
    fi
    $M := \{C \mid C$ is a predecessor of a $D \in M\}$
end.

Now, to any $L \in C$, where C is a strong component, we assign the truth value of C to L. This gives us a truth value assignment satisfying G.

(b) $\Rightarrow$ (c) If G contains a contradictory strong component, we have a pair of literals $L, \bar{L}$ and two paths U and V in G, such that $P=L \to U \to \bar{L}$ and $Q=\bar{L} \to V \to L$ are paths of G. If $\mathrm{Lit}(U) \cap \mathrm{Lit}(V)=\varnothing$, we have a contradictory cycle in G. Otherwise, let K be the first literal of U common to U and V. Then $P=L \to U_1 \to K \to U_2 \to \bar{L}$ and $Q=\bar{L} \to V_1 \to K \to V_2 \to L$ and $\mathrm{Lit}(U_1) \cap \mathrm{Lit}(V)=\varnothing$. If $\mathrm{Lit}(U_1) \cap \mathrm{Lit}(\bar{V}_2) \neq \varnothing$ the cycle in Fig. 2 is a contradictory cycle of G.

If $\mathrm{Lit}(U_2) \cap \mathrm{Lit}(\bar{V}_1) \neq \varnothing$ the two paths $\bar{L} \to V_1 \to K$, $K \to U_2 \to \bar{L}$ give us two paths like P and Q above, with the number of common literals reduced by at least 1 and we can proceed by induction. Finally, let $\mathrm{Lit}(U_2) \cap \mathrm{Lit}(\bar{V}_1)=\varnothing$ and $\mathrm{Lit}(U_1) \cap \mathrm{Lit}(\bar{V}_2)=\varnothing$. Let I be the last literal on $L \to U_1$ with $\bar{I} \in \mathrm{Lit}(\bar{L} \to V_1)$ and let U_1' be the piece of U_1 to the right of I and V_1' the piece of V_1 to the right of $\bar{I}$. Let J be the first literal on $U_2 \to \bar{L}$ with $\bar{J} \in \mathrm{Lit}(V_2 \to L)$. Again let U_2' be the piece of U_2 to the left of J and V_2' the piece of V_2 to the left of $\bar{J}$. Then the two paths $I \to U_1' \to K \to U_2' \to J$ and $J \to \bar{V}_2' \to \bar{K} \to \bar{V}_1' \to I$ contain no common literal and thus can be connected to a contradictory cycle.

(c) $\Rightarrow$ (a) (cf. also [22]) The edges of the formula graph present the implications whose conjunction is satisfiable iff the formula is satisfiable. If the formula graph has a contradictory cycle containing the literals L and $\bar{L}$ we cannot satisfy the formula: If we assign a 1 to L, $\bar{L}$ must be 1, too, in order to satisfy all implications. If we assign 1 to $\bar{L}$, L must be 1. In either case we cannot satisfy all implications. ∎

We need the following arithmetical formulas: Let $p \geqslant q \geqslant 0$. Then

$$(p)_q=\prod_{i=1}^{q}(p-i+1)=p \cdot (p-1) \cdot \cdots \cdot (p-q+1)$$

and

$$\binom{p}{q}=\frac{(p)_q}{q!}.$$

$$\pi = \begin{array}{ccccccccccccccc} S_1 & \to & U_1 & \to & S_2 & \to & U_2 & \to & \dots & & \to & S_a & \to & U_a \\ \uparrow & & & & & & & & & & & & & \downarrow \\ V_a & \leftarrow & \overline{S}_a & \leftarrow & & & & \dots & \leftarrow & V_2 & \leftarrow & \overline{S}_2 & \leftarrow & V_1 & \leftarrow & \overline{S}_1 \end{array}$$

FIGURE 3

If $s \leqslant q$ then

$$\frac{\binom{p-s}{q-s}}{\binom{p}{q}} = \frac{(q)_s}{(p)_s}.$$

If $p = p(n)$, $q = q(n)$, and $q = o(p^{1/2})$ then $(p)_q \sim p^q$; see [14, formula 3.5, p. 130].

2. FEWER CLAUSES THAN VARIABLES

In this section we prove:

2.1. Theorem. *If* $C < 1$ *and* $q(n) \leqslant C \cdot n$ *then almost all formulas of* $\text{Form}_n(q(n))$ *are satisfiable. That is* $\Pr(\text{Sat}(q(n))) \to 1$.

The proof of Theorem 2.1 is presented as a series of definitions and lemmas. For the rest of this section let:

$C = 1 - \varepsilon$, where $\varepsilon > 0$ is fixed;

$q = q(n) \in \mathcal{N}$ with $q(n) \leqslant C \cdot n$.

If not otherwise stated, the results of this section refer to the probability spaces $\text{Form}_n(q(n))$ and $FG_n(q(n))$.

We first show that a contradictory formula graph, i.e., a formula graph with a contradictory cycle, has a contradictory cycle in a certain normalform.

2.2. Definition. Let $a, b, c \in \mathcal{N}$ with $a \geqslant 1$, $b \geqslant a$, and $c \geqslant 0$. The cycle π is of type a, b, c iff π can be decomposed as in Fig. 3, where the S_i are nonempty paths and the U_i, V_i are (possibly empty) paths such that:

- Each contradictory pair of π occurs in exactly one S_i, $\overline{S}_i$.
- $\sum_i |\text{Lit}(S_i)| = b$ ($=$ the number of contradictary pairs of π).
- $\sum_i (|\text{Lit}(U_i)| + |\text{Lit}(V_i)|) = c$ ($=$ the number of literals $L \in \pi$ with $\overline{L} \notin \pi$).

$$\begin{array}{ccc} x & \to & y \\ \uparrow & & \downarrow \\ \overline{y} & \to & \overline{x} \end{array}$$

FIGURE 4

$$\begin{array}{ccccc} x & \longrightarrow & y & \longrightarrow & u \\ \uparrow & & & & \uparrow \\ v & \longleftarrow & \overline{x} & \longleftarrow & \overline{y} \end{array}$$

FIGURE 5

We say π has a sections. The number of pairs of edges to get π is given by: $|\text{Pairs}(\pi)| = a + b + c$, whereas $\text{Length}(\pi) = a + b + c + (b - a)$. Hence, $b - a$ is the number of pairs of complementary edges of π. We say π is in normalform if there exist a, b, c such that π is of the type a, b, c.

2.3. Example. The cycle of Fig. 4 is of type 2, 2, 0 with $S_1 = x$ and $S_2 = y$. Note that we cannot have $S_1 = x \to y$ because then $\overline{S}_1 = \bar{y} \to \bar{x}$.

The cycle of Fig. 5 is of type 1, 2, 2 with $S_1 = x \to y$, $U_1 = u$, and $V_1 = v$.

The cycle of Fig 6 is not in normalform. The complementary literals are $x, \bar{x}, y, \bar{y}$. If $S_1 = x$, then $S_2 = y$ but then we should have $\bar{y}$ after $\bar{x}$. If $S_1 = x \to u \to y$ then $\overline{S}_1 = \bar{y} \to \bar{u} \to \bar{x}$ does not occur in the cycle. But a formula graph containing this cycle also contains the path $\bar{y} \to \bar{u} \to \bar{x}$ and therefore the cycle in Fig. 7 which is of type 1, 3, 2 with $S_1 = x \to u \to y$, three pairs of complementary literals, two literals whose complement is not on the cycle.

Note that the decomposition of a cycle is unique, up to cyclic permutations of the $S_1, ..., S_a, \overline{S_1}, ..., \overline{S_a}$.

2.4. Lemma. *Every contradictory formula graph has a cycle in normalform.*

Proof. Let π be a contradictory cycle in the formula graph G. If π has only one contradictory pair, π is in normalform. If π has more contradictory pairs we show in two steps that G has a cycle in normalform. In Step 1 we show that G has a contradictory cycle which can be decomposed as $\circlearrowright^{V}_{W}$ such that neither V nor W alone contain a contradictory pair of literals. In Step 2 we show that if G has a cycle which can be decomposed in V and W as above then G has a cycle in normalform.

Step 1. We decompose π as $\pi = \circlearrowright^{V}_{W}$ where V, W is a pair of paths such that V contains the minimal number of

$$\begin{array}{ccccccc} x & \to & u & \to & y & \to & v \\ \uparrow & & & & & & \downarrow \\ s & \leftarrow & \overline{x} & \leftarrow & r & \leftarrow & \overline{y} \end{array}$$

FIGURE 6

$$\begin{array}{ccccccc} x & \rightarrow & u & \rightarrow & y & \rightarrow & v \\ \uparrow & & & & & & \downarrow \\ s & \leftarrow & \bar{x} & \leftarrow & \bar{u} & \leftarrow & \bar{y} \end{array}$$

FIGURE 7

contradictory pairs among all pairs of paths R, S with $\pi = \circlearrowright\!{R \atop S}\!\circlearrowleft$ satisfying:

(1) S contains no contradictory pair of literals.

(2) S contains at least one literal L with $\bar{L} \in \pi$.

As two paths R, S, where $\pi = \circlearrowright\!{R \atop S}\!\circlearrowleft$ and $S = L$ is a single literal with $\bar{L} \in \pi$, satisfy (1) and (2), such a decomposition exits.

We proceed inductively on the number of contradictory pairs in V. If V contains no such pair we are finished. Otherwise, let $L, \bar{L} \in V$ be the unique pair such that $\bar{L}$ follows L on V and such that for each $F \in V$ following $\bar{L}$ we have $\bar{F} \notin V$. Let H be the first literal after $\bar{L}$ such that $\bar{H} \in W$. H must exist because of the minimality condition for V. Otherwise $\bar{L}$ would belong to W. We decompose π as in Fig. 8 for suitable paths X_i. The formula graph G contains the path $\bar{H} \rightarrow \bar{X}_2 \rightarrow L$. For all $K \in X_2$ holds $\bar{K} \notin \pi$. This follows from the definition of $\bar{L}$ and H. Therefore we get a cycle π' if we replace $\bar{H} \rightarrow X_4 \rightarrow L$ in π by $\bar{H} \rightarrow \bar{X}_2 \rightarrow L$. We can decompose $\pi' = \circlearrowright\!{V' \atop W'}\!\circlearrowleft$ such that V', W' satisfy (1) and (2) above but V' contains at least one contradictory pair (the pair $L, \bar{L}$) less that V. The claim follows by induction.

Step 2. Let π be a cycle which has been obtained by Step 1 and has at least 2 contradictory pairs. We decompose π as $\pi = \circlearrowright\!{V \atop W}\!\circlearrowleft$ such that $L, \bar{L} \in \pi \Rightarrow L \in V, \bar{L} \in W$.

Let $S_1, ..., S_\alpha$ where $\alpha \geqslant 1$ be the set of maximal nonempty subpaths of V such that $\bar{S}_1, ..., \bar{S}_\alpha$ are subpaths of W. Then we have paths $U_1, ..., U_\alpha, V_1, ..., V_\alpha$ such that π can be decomposed as Fig. 9, where $\{i_1, ..., i_\alpha\} = \{1, ..., \alpha\}$. If for all k, l holds $l > k \Rightarrow i_l > i_k$, the cycle is in normalform. Otherwise, we proceed inductively on the number of pairs l, k violating this condition. Let h be the smallest k for which there is an $l > k$ with $i_l < i_k$. Then the lower path of π has the form of Fig. 10 and $i_h > h$.

Let j be the greatest l with $l > h$ and $i_l < i_h$. We replace the path $\bar{S}_{i_j} \leftarrow \cdots \leftarrow \bar{S}_{i_h}$ in the lower half of π by the complement of $S_{i_j} \rightarrow U_{i_j} \rightarrow \cdots \rightarrow U_{i_{h-1}} \rightarrow S_{i_h}$ from the upper half. This replacement gives us a cycle because we have $i_j \geqslant h$ and

$$\pi = \begin{array}{ccccccccc} L & \rightarrow & X_1 & \rightarrow & \bar{L} & \rightarrow & X_2 & \rightarrow & H \\ & \nwarrow & & & & & & \swarrow & \\ & & X_4 & \leftarrow & \bar{H} & \leftarrow & X_3 & & \end{array}$$

FIGURE 8

$$\pi = \begin{array}{cccccccccccccc} S_1 & \rightarrow & U_1 & \rightarrow & S_2 & \rightarrow & U_2 & \rightarrow & \ldots & & \rightarrow & S_\alpha & \rightarrow & U_\alpha \\ \uparrow & & & & & & & & & & & & & \downarrow \\ V_\alpha & \leftarrow & \bar{S}_{i_\alpha} & \leftarrow & & & \ldots & \leftarrow & V_2 & \leftarrow & \bar{S}_{i_2} & \leftarrow & V_1 & \leftarrow \; \bar{S}_{i_1} \end{array}$$

FIGURE 9

$l > j \Rightarrow i_l > i_h$. Moreover, the number of pairs violating the condition above has decreased by 1. ∎

2.5. Definition. Let π be a cycle in normalform. X_π is the indicator random variable of the event "G contains π." That is,

$$X_\pi \colon FG(q) \rightarrow \{0, 1\}$$

with

$$X_\pi(G) = \begin{cases} 1 & \text{if } G \text{ contains } \pi, \\ 0 & \text{otherwise.} \end{cases}$$

We have $X_\pi = 0$ if π is of type a, b, c with $a + b + c > q$. In the following we always assume $a + b + c \leqslant q$.

The random variable $X_{a,b,c}$ is given by

$$X_{a,b,c} = \sum_\pi X_\pi,$$

where the sum goes over all cycles of type a, b, c in the complete formula graph over n variables. The random variable X is given by

$$X = \sum_{a,b,c} X_{a,b,c},$$

where the sum goes over all a, b, c with $a \geqslant 1$, $b \geqslant a$, and $c \geqslant 0$. $X(G)$ gives us the number of all cycles in normalform in G. Let $\mu_{a,b,c}$ be the number of all cycles of type a, b, c in the complete formula graph.

To prove Theorem 2.1 we now proceed as follows: We show $\Pr(\mathrm{Unsat}(q)) \rightarrow 0$. We have

$$\begin{aligned} \Pr(\mathrm{Unsat}(q)) &= \Pr\{G \in FG(q) \mid G \text{ contains} \\ &\qquad \text{a contradictory cycle}\} \\ &= \Pr(X \geqslant 1) \leqslant EX, \end{aligned}$$

where that last inequality is Markov's inequality.

We show $EX \rightarrow 0$ for $n \rightarrow \infty$.

$$\ldots \leftarrow \bar{S}_{i_h} \leftarrow E_{h-1} \leftarrow \bar{S}_{i_{h-1}} \leftarrow \ldots \leftarrow E_1 \leftarrow \bar{S}_1$$

FIGURE 10

2.6. COROLLARY. (a) *Let π be a cycle of type a, b, c, then*

$$EX_\pi = \frac{(q)_{a+b+c}}{(N)_{a+b+c}}$$

(b) $$\mu_{a,b,c} \leqslant \binom{n}{b} * b! * 2^b * \binom{n-b}{c} * c! * 2^c * \binom{b-1}{a-1} * \binom{c+2a-1}{c} * \frac{1}{2a}$$

(c) $$EX_{a,b,c} \leqslant \binom{b-1}{a-1} * \binom{c+2a-1}{2a-1} * C^{b+c} * \left(\frac{D}{n}\right)^a,$$

$$\text{where} \quad D = \frac{C}{1-C}.$$

(*Recall*: *C is defined right after Theorem* 2.1.)

Proof.

(a) $$EX_\pi = \sum_{G \in FG(q)} X_\pi(G) * \Pr(G) = \frac{\{G \in FG(q) \mid G \text{ contains } \pi\}}{\binom{N}{q}} = \frac{\binom{N-a-b-c}{q-a-b-c}}{\binom{N}{q}} = \frac{(q)_{a+b+c}}{(N)_{a+b+c}}$$

$$\text{as} \quad a+b+c \leqslant q \leqslant N.$$

Note that π requires $a+b+c$ complementary pairs of edges.

(b) Let

$$B = \binom{n}{b} * b! * 2^b, \qquad E = \binom{n-b}{c} * c! * 2^c.$$

Each cycle π of type a, b, c can be obtained as follows:

(1) Choose the path $S = S_1 \to S_2 \to \cdots \to S_a$: B possibilities. We choose b variables from n order them and negate them or not. Note that we only choose the whole path S not its partition into the S_i.

(2) Choose the path $U = U_1 \to U_2 \to \cdots \to U_a \to V_1 \to \cdots \to V_a$: E possibilities. Again we do not yet choose the partition of U.

(3) Partition S into the S_i where each S_i is nonempty: $\binom{b-1}{a-1}$ possibilities. The number of choices for the S_i is equal to the number of vectors $(m_1, ..., m_a)$ with $m_i \in \mathcal{N} \setminus \{0\}$ for all i and $m_1 + \cdots + m_a = b$. There are $\binom{b-1}{a-1}$ such vectors [15, Exercise 11, p. 13].

(4) Partition the path U into the U_i, V_i: $\leqslant \binom{c+2a-1}{c}$ possibilities: The number of choices for the U_i, V_i is bounded by the number of vectors $(m_1, ..., m_{2a})$ with $m_i \in N$ ($m_i = 0$ is possible) and $m_1 + \cdots + m_{2a} = c$. There are $\binom{c+2a-1}{c} = \binom{c+2a-1}{2a-1}$ such vectors [15, Proposition 6.1, p. 11]. (In the case $a=1$ we have $U_1 \neq \varnothing$ and $V_1 \neq \varnothing$ because we have no edges $x \to \bar{x}$, therefore the $\leqslant$.)

(1) through (4) give us

$$\leqslant B * E * \binom{b-1}{a-1} * \binom{c+2a-1}{2a-1}$$

choices for π. But, each π has been chosen $2 * a$ times because of cyclic permutations and because choosing $S = S_1 \to \cdots \to S_a$ and $S = \bar{S}_1 \to \cdots \to \bar{S}_a$ gives the same π twice. So the above product divided by $2a$ is still an upper bound for $\mu_{a,b,c}$.

(c) Some simple bounds first:

$$(q(n))_{a+b+c} \leqslant (q(n))^{a+b+c} \leqslant C^{a+b+c} * n^{a+b+c};$$

$$\begin{aligned}
(N)_{a+b+c} &= \prod_{i=1}^{a+b+c} (N-i+1) \\
&\geqslant \prod (2n(n-1) - 2(i-1)) \\
&\qquad (\text{as } 2(a+b+c-1) \leqslant 2n(n-1)) \\
&= 2^{a+b+c} * \prod (n(n-1) - i + 1) \\
&\geqslant 2^{a+b+c} * \prod (n(n-1) - n(i-1)) \\
&\qquad (\text{as } a+b+c-1 \leqslant n-1) \\
&= 2^{a+b+c} * n^{a+b+c} * (n-1)_{a+b+c};
\end{aligned}$$

$$\begin{aligned}
(n-b-c)_{a+1} &= \prod_{i=1}^{a+1} (n-b-c-i+1) \\
&\geqslant (n-q(n))^{a+1} \\
&\qquad (\text{as } a+b+c \leqslant q(n) \leqslant n) \\
&\geqslant (n - C * n)^{a+1} \\
&\qquad (\text{as } q(n) \leqslant C * n \leqslant n).
\end{aligned}$$

Finally,

$$\begin{aligned}
EX_{a,b,c} &\leqslant \mu_{a,b,c} * \frac{(q)_{a+b+c}}{(N)_{a+b+c}} \\
&\leqslant (n)_{b+c} * 2^{b+c} * \binom{b-1}{a-1} * \binom{c+2a-1}{2a-1} * \frac{1}{2a} \\
&\quad * \frac{C^{a+b+c} * n^{a+b+c}}{2^{a+b+c} * n^{a+b+c} * (n-1)_{a+b+c}}
\end{aligned}$$

$$\leqslant \binom{b-1}{a-1} * \binom{c+2a-1}{2a-1} * \frac{1}{2a} * \frac{1}{2^a} * C^{a+b+c} * \frac{n}{(n-C*n)^{a+1}}$$
$$= \binom{b-1}{a-1} * \binom{c+2a-1}{2a-1} * \frac{1}{2a} * \frac{1}{2^a} * C^{b+c} * \frac{1}{1-C} * \left(\frac{C}{(1-C)*n}\right)^a$$
$$\leqslant \binom{b-1}{a-1} * \binom{c+2a-1}{2a-1} * C^{b+c} * \left(\frac{D}{n}\right)^a,$$

where $D = C/(1-C)$. ∎

The following theorem finishes the proof of Theorem 2.1.

2.7. THEOREM. *$EX \to 0$ for $n \to \infty$.*

Proof. With Corollary 2.6(c) we get as $C < 1$,

$$EX \leqslant \sum_{a=1}^{\infty} \left(\frac{D}{n}\right)^a * \sum_{b-1=a-1}^{\infty} \binom{b-1}{a-1} * C^b * \sum_{c=0}^{\infty} \binom{c+2a-1}{2a-1} * C^c$$
$$= \sum_{a=1}^{\infty} \left(\frac{D}{n}\right)^a * C * \sum_{b-1=a-1}^{\infty} \binom{b-1}{a-1} * C^{b-1} * \frac{1}{(1-C)^{2a}} \qquad [11, \text{formula } 5.56]$$
$$= \sum_{a=1}^{\infty} \left(\frac{D}{n}\right)^a * \frac{C^a}{(1-C)^a} * \frac{1}{(1-C)^{2a}} \qquad [11, \text{formula } 5.57]$$
$$= \sum_{a=1}^{\infty} \left(\frac{E}{n}\right)^a = \frac{1}{1-E/n} - 1 \to 0 \qquad \text{for} \quad n \to \infty,$$

where E is a suitable constant. ∎

3. MORE CLAUSES THAN VARIABLES

In this section we prove:

3.1. THEOREM. *If $C > 1$ and $q(n) \in \mathcal{N}$ with $q(n) \geqslant C \cdot n$ then almost all formulas of* $\mathrm{Form}_n(q(n))$ *are unsatisfiable. That is* $\Pr\{F \in \mathrm{Form}_n(q(n)) \mid F \textit{ is unsatisfiable}\} \to 1$ *for $n \to \infty$.*

For the rest of this section let n be large and:

$$N = N(n) = 4 \cdot \binom{n}{2},$$
$$C = 1 + \varepsilon, \qquad \text{where} \quad \varepsilon > 0 \text{ is fixed},$$
$$q = q(n) \in \mathcal{N} \qquad \text{with} \quad N(n) \geqslant q(n) \geqslant C \cdot n,$$
$$k = k(n) \in \mathcal{N} \qquad \text{with} \quad k(n) = \lceil \log_C n \rceil,$$
$$l = l(n) = 2k + 2.$$

We always refer to the probability spaces $\mathrm{Form}_n(q(n))$ and $FG_n(q(n))$.

For a, b, c fixed we get with Corollary 2.6(c) that if $q(n) = C \cdot n$ then $EX_{a,b,c} \leqslant O((1/n)^a)$. As $a \geqslant 1$ almost no graph has a cycle of type a, b, c for fixed a, b, c. Since the transformation proving Lemma 2.4 transforms contradictory cycles of length $O(1)$ into cycles of type a, b, c, where $a + b + c = O(1)$, we also get that almost no graph has a contradictory cycle of length $O(1)$. Therefore we look at cycles of length logarithmic in n, actually of length $l = 2k + 2$. We show that almost all formula graphs have a cycle π of length l, such that π contains exactly one contradictory pair of literals. Intuitively, cycles with more contradictory pairs only seem less likely for the following reason: For a cycle of length l with $m \geqslant 1$ contradictory pairs we can choose only $l - m$ vertices because m vertices participate in the contradiction.

At this point it might be useful to look at a somewhat similar situation in normal directed random graphs: The formula graph of a random formula $\mathrm{Form}_n(C \cdot n)$ is a directed random graph over $2 \cdot n$ vertices with $C \cdot n$ independent *pairs* of edges, hence with $2 \cdot C \cdot n$ edges. If we look at the analogous situation of a directed random graph with $2 \cdot n$ vertices and $C \cdot 2 \cdot n$ independent edges we have [13]: If $C < 1$, then the probability that such a graph has a cycle approaches C and the average number of cycles $-\log(1-C)$. As the probability C, approaches 1 as $C \to 1$, for $C \geqslant 1$ almost all directed random graphs have a cycle. (Here one uses the trivial observation that more edges are more likely to induce a cycle.)

In the next definition we fix the kind of cycles which we consider.

3.2. DEFINITION. (a) Let P be a (nonempty) path and L a literal. We consider the unordered pair of paths $L \to P \to \bar{L}$, $L \to \bar{P} \to \bar{L} (= \overline{L \to P \to \bar{L}})$. If a formula graph has one of these paths it has the other one, too. We distinguish these two paths by calling one of them *main path* and the other one *side path*. We assume that for any pair of paths as above these names are fixed from now on.

(b) A cycle π is a *simple cycle* iff there exist a variable x and main paths

$$x \to L_1 \to L_2 \to \cdots \to L_k \to \bar{x},$$
$$\bar{x} \to G_1 \to G_2 \to \cdots \to G_k \to x$$

and such that π can be presented as in Fig.11, where $x, \bar{x}$ is the only contradictory pair of π (and all nodes are pairwise different). Note that π contains no pair of complementary edges. We call $x \to L_1 \to L_2 \to \cdots \to L_k \to \bar{x}$ the *first main path* of π, Fmp π, and $\bar{x} \to G_1 \to G_2 \to \cdots \to G_k \to x$ the *second main path* of π, Smp π. The *first* and *second side path* of π are given by Fsp $\pi = \overline{\text{Fmp}\,\pi}$ and Ssp $\pi = \overline{\text{Smp}\,\pi}$. We

$$\pi = \begin{array}{ccccccccccc} x & \to & L_1 & \to & L_2 & \to & \dots & & & & \to L_k \\ \uparrow & & & & & & & & & & \downarrow \\ G_k & \leftarrow & & & \dots & \leftarrow & G_2 & \leftarrow & G_1 & \leftarrow & \bar{x} \end{array}$$

FIGURE 11

have $|\mathrm{Pairs}(\mathrm{Fmp}\,\pi)| = |\mathrm{Pairs}(Smp)\,\pi| = k+1$ and $|\mathrm{Pairs}(\pi)| = l$. Note that $\mathrm{Fsp}\,\pi = x \to \bar{L}_k \to \bar{L}_{k-1} \to \cdots \to \bar{L}_1 \to \bar{x}$ and $\mathrm{Ssp}\,\pi = \bar{x} \to \bar{G}_k \to \bar{G}_{k-1} \to \cdots \to \bar{G}_1 \to x$. The first paths always go from a variable x to $\bar{x}$, whereas the second paths go from a negated variable $\bar{x}$ to x.

(c) Let π be a simple cycle in the complete formula graph over n variables. X_π is the indicator random variables of the event "G contains π":

$$X_\pi = \begin{cases} 1 & \text{if} \quad G \text{ contains } \pi \\ 0 & \text{otherwise.} \end{cases}$$

The random variables X gives us the total number of simple cycles:

$$X = \sum_\pi X_\pi,$$

where the sum is over all simple cycles π in the complete formula graph. Let $\mu = \mu(n)$ be the total number of simple cycles in the complete formula graph over n variables.

3.3. COROLLARY.

(a) $$\mu = n * \binom{n-1}{k} * k! * 2^{k-1} * \binom{n-1-k}{k} * k! * 2^{k-1} = (n)_{l-1} * 2^{2k-2} \sim n^{l-1} * 2^{l-4}.$$

(b) $$EX_\pi = \frac{\binom{N-l}{q-l}}{\binom{N}{q}} = \frac{(q)_l}{(N)_l} = \left(\frac{C}{2*(n-1)}\right)^l * (1+o(1)).$$

(c) $$EX = \mu * \frac{\binom{N-l}{q-l}}{\binom{N}{q}} = \mu * \frac{(q)_l}{(N)_l} \geqslant \mu * \left(\frac{C}{2*(n-1)}\right)^l \geqslant \frac{1}{16} * C^2 * n * (1+o(1)) \to \infty$$

as $\quad n \to \infty$.

Proof. (a) Each simple cycle in the complete formula graph can be obtained by exactly one sequence of choices of the following choosing process:

(1) Choose the contradictory pair: n possibilities.

(2) Choose the k variables of the first main path, order them, and negate them or not: $\binom{n-1}{k} * k! * 2^{k-1}$ possibilities. The last factor is 2^{k-1} instead of 2^k because if $x \to P \to \bar{x}$ is a main path, then $x \to \bar{P} \to \bar{x}$ is *not* a main path and therefore cannot be chosen.

(3) The second main path: $\binom{n-1-k}{k} * k! * 2^{k-1}$ possibilities.

Hence for n large enough

$$\begin{aligned} \mu &= n * \binom{n-1}{k} * k! * 2^{k-1} * \binom{n-1-k}{k} * k! * 2^{k-1} \\ &= n * \frac{(n-1)!}{k!(n-1-k)!} * k! * 2^{k-1} * \frac{(n-1-k)!}{k!(n-1-2k)!} \\ &\quad * k! * 2^{k-1} \\ &= (n)_{2k+1} * 2^{2k-2} \sim n^{l-1} * 2^{l-4} \end{aligned}$$

because $l = o(n^{1/2})$ and $l = 2k+2$.

(b) Let π be a simple cycle in the complete formula graph. There are $\binom{N-l}{q-l}$ graphs in $FG(q)$ which have π because $|\mathrm{Pairs}(\pi)| = l$; hence

$$\begin{aligned} EX_\pi &= \frac{1}{\binom{N}{q}} * \sum_G X_\pi(G) = \frac{\binom{N-l}{q-l}}{\binom{N}{q}} = \frac{(q)_l}{(N)_l} \\ &\geqslant \frac{(C*n)_l}{(2*n*(n-1))_l} = \left(\frac{C}{2(n-1)}\right)^l * (1+o(1)) \end{aligned}$$

because $l = o(n^{1/2})$.

(c) $$\begin{aligned} EX &= \sum_\pi EX_\pi = \sum_\pi \frac{\binom{N-l}{q-l}}{\binom{N}{q}} = \mu * \frac{\binom{N-l}{q-l}}{\binom{N}{q}} = \mu * \frac{(q)_l}{(N)_l} \\ &\geqslant n^{l-1} * 2^{l-4} * \frac{C^{2k} * C^2}{2^l * (n-1)^l} * (1+o(1)) \\ &\geqslant \frac{1}{16} C^2 * n * (1+o(1)), \end{aligned}$$

as $l = o(n^{1/2})$ and $k = \lceil \log_C n \rceil$. Here we need $C > 1$! ∎

In the following we use the abbreviations

$$\mu_1 = \binom{n-1}{k} * k! * 2^{k-1} = (n-1)^k * 2^{k-1} * (1+o(1)),$$

$$\mu_2 = \binom{n-1-k}{k} * k! * 2^{k-1}$$

$$= (n-1-k)^k * 2^{k-1} * (1+o(1)),$$

as $k = o(n^{1/2})$. Then $\mu = n * \mu_1 * \mu_2$.

Despite the fact that the average number of simple cycles goes to infinity we cannot be sure that each graph has a simple cycle. To prove this we apply the second moment method extending the argument of [16]: We have

$$\Pr\{F \mid F \text{ is satisfiable}\}$$
$$\leqslant \Pr\{G \in FG(q) \mid G \text{ has no simple cycle}\}$$
$$= \Pr(X=0)$$
$$\leqslant \Pr((X-EX)^2 \geqslant (EX)^2) \leqslant \frac{VX}{(EX)^2},$$

where the last inequality is *Chebychef's inequality* [15, p. 249]. We show

$$\frac{VX}{(EX)^2} = o(1).$$

As

$$\frac{VX}{(EX)^2} = \frac{E(X^2)}{(EX)^2} - 1,$$

the result follows from

$$\frac{E(X^2)}{(EX)^2} = 1 + o(1).$$

As

$$X = \sum_{\pi} X_\pi,$$

we have

$$X^2 = \sum_{(\pi, \pi')} X_\pi * X_{\pi'},$$

where the sum is over all ordered pairs of simple cycles (π, π') in the complete formula graph. To compute $E(X^2)/(EX)^2$ we decompose the random variable X^2 as a sum of random variables. This decomposition is induced by the different ways in which the two simple cycles (π, π') can have common edges.

3.4. Definition. (a) Let V be a path in a formula graph, let $\varnothing \neq S \subseteq E(V)$ ($E(V)$ is the set of edges of V) and let $C = e_1 \cdots e_n$, where $n \geqslant 1$ be a sequence of n consecutive edges of V.

C is a *chain* of S on V iff $C \in S$ and if f is the edge succeding e_n or preceding e_1 on V then $f \notin S$.

We let $\text{Chain}_V S$ be the set of chains of S on V. We have $1 \leqslant |\text{Chain}_V S| \leqslant |S|$ and $|\text{Chain}_V S| \leqslant \text{Length}(V) - |S| + 1$. We need the "$+1$" because if $V = A \to B \to C \to D$ and $S = \{A \to B, C \to D\}$, then there are two chains of S on V and $\text{Length}(V) - |S| + 1 = 2$. If $S = \{A \to B, B \to C\}$, we have one chain of S on V.

(b) For an ordered pair (π, π') of simple cycles we define conditions 1 through 7. By these conditions we distinguish the ways in which $\pi \cup \bar{\pi}$ and $\pi' \cup \bar{\pi}'$ can have common edges.

1. $\pi = \pi'$.

The remaining conditions only apply if $\pi \neq \pi'$. Let $E = E(\pi \cup \bar{\pi})$ be the set of edges of $\pi \cup \bar{\pi}$.

2. $E(\pi') \cap E = \varnothing$. $\pi \cup \bar{\pi}$ and π' have no common edges (they may have common vertices).

3. $\text{Fmp}\,\pi' = \text{Fmp}\,\pi$ and
   a. $E(\text{Smp}\,\pi') \cap E = \varnothing$ or
   b. $E(\text{Smp}\,\pi') \cap E = S \neq \varnothing$.

We say the pair (π, π') satisfies conditions 3.b with parameters r, R iff $|S| = r$ and $|\text{Chain}_{\text{Smp}\,\pi'} S| = R$. (Then $1 \leqslant r \leqslant k$, $1 \leqslant R \leqslant r$, and $R \leqslant k + 1 - r + 1 = k - r + 2$.)

4. a, b. Analogous to 3, changing the roles of Fmp and Smp.

5. $E(\text{Fmp}\,\pi') \cap E = \varnothing$, $\text{Smp}\,\pi' \neq \text{Smp}\,\pi$, and $E(\text{Smp}\,\pi') \cap E = S \neq \varnothing$. As in 3.b we define condition 5 with parameters r, R.

6. Analogous to 5, changing to roles of Smp and Fmp.

7. $\text{Fmp}\,\pi' \cap E = S \neq \varnothing$, $\text{Fmp}\,\pi' \neq \text{Fmp}\,\pi$, and $\text{Smp}\,\pi' \cap E = S' \neq \varnothing$, $\text{Smp}\,\pi' \neq \text{Smp}\,\pi$. We say (π, π') satisfies condition 7 with parameters r, R and t, T iff $|S| = r$, $|\text{Chain}_{\text{Fmp}\,\pi'} S| = R$ and $|S'| = t$, $|\text{Chain}_{\text{Smp}\,\pi'} S'| = T$.

For each pair of simple cycles exactly one of the above conditions holds. Note that situations like $\text{Fmp}\,\pi' = \text{Fsp}\,\pi$ cannot occur because $\text{Fmp}\,\pi'$ must be a main path. Situations like $\text{Fmp}\,\pi' = \text{Smp}\,\pi$ cannot occur because first main paths go from a positive to a negative variable, whereas second main paths go from a negative to a positive variable.

(c) The random variable X_1 is given by

$$X_1 : FG(q) \to \mathcal{N}$$

$$X_1 = \sum_{\pi, \pi'} X_\pi * X_{\pi'},$$

where the sum goes over all pairs of simple cycles satisfying condition 1. The random variables X_2, X_{3a}, X_{3b}, X_{4a}, X_{4b}, X_5, X_6, and X_7 are defined analogously.

Now we have

$$X^2 = X_1 + X_2 + \cdots + X_7$$

and

$$\frac{E(X^2)}{(EX)^2} = \frac{EX_1}{(EX)^2} + \frac{EX_2}{(EX)^2} + \cdots + \frac{EX_7}{(EX)^2}.$$

In the followings series of lemmas we compute these summands. It turns out that

$$\frac{EX_j}{(EX)^2} = o(1), \qquad \text{i.e., } EX_j = o((EX)^2)$$

for $j \neq 2$, whereas

$$\frac{EX_2}{(EX)^2} \sim 1, \qquad \text{i.e., } EX_2 \sim (EX)^2.$$

As X_2 counts the number of pairs of simple cycles without common edges, this means that the average number of pairs of simple cycles (π, π') such that π' has common edges with $\pi \cup \bar{\pi}$ is asymptotically irrelevant, compared to the number of such pairs without common edges. An intuitive explanation why this property implies that almost all graphs have a simple cycle is the following: Assume we have a constant fraction of all formula graphs contains no simple cycle; then the remaining formula graphs must be packed very densely with simple cycles in order to allow for the high average. But, the more densely the graphs are packed with cycles, the more often pairs of cycles having edges in common occur.

3.5. Lemma. $EX_1/(EX)^2 = o(1)$.

Proof. Trivial, as $X_1 = X$ and $EX \to \infty$. ∎

3.6. Lemma. $EX_2/(EX)^2 \leqslant 1$.

Proof. The number of pairs (π, π') with $E(\pi') \cap E(\pi \cup \bar{\pi}) = \varnothing$ is $\leqslant \mu^2$. For a given (π, π') we have $\binom{N-2l}{q-2l}$ graphs in $FG(q)$ which have both π and π'. Hence

$$EX_2 \leqslant \mu^2 * \binom{N-2l}{q-2l} \Big/ \binom{N}{q}$$

and with Corollary 3.3(c)

$$\frac{EX_2}{(EX)^2} \leqslant \frac{\binom{N-2l}{q-2l}}{\binom{N-l}{q-l}} * \frac{\binom{N}{q}}{\binom{N-l}{q-l}}$$

$$= \frac{(q-l)_l}{(N-l)_l} * \frac{(N)_l}{(q)_l} \leqslant \frac{(q)_l}{(N)_l} * \frac{(N)_l}{(q)_l} = 1,$$

where the last inequality holds because $(q-l-i)/(N-l-i) \leqslant (q-i)/(N-i)$ which follows from the trivial observation that for $u, v, w > 0$ holds $u \leqslant v \Rightarrow u/v \leqslant (u+w)/(v+w)$. ∎

3.7. Lemma. $EX_{3a}/(EX)^2 = o(1)$, $EX_{4a}/(EX)^2 = o(1)$.

Proof. For fixed π we have $\leqslant \mu_2$ many cycles π' such that (π, π') satisfies condition 3.a because $\mathrm{Fmp}\,\pi'$ (and hence the contradictory pair of π) cannot be chosen. Hence, the number of pairs of cycles satisfying 3.a is $\leqslant \mu * \mu_2$. For a given (π, π') satisfying 3.a we have $\binom{n-2l+k+1}{q-2l+k+1}$ graphs in $FG(q)$ containing both π and π'. Hence,

$$EX_{3a} \leqslant \mu * \mu_2 * \binom{N-2l+k+1}{q-2l+k+1} \Big/ \binom{N}{q}$$

and with corollary 3.3(c),

$$\frac{EX_{3a}}{(EX)^2} \leqslant \frac{\binom{N-l-k-1}{q-l-k-1}}{\binom{N-l}{q-l}} * \frac{\binom{N}{q}}{\binom{N-l}{q-l}} * \frac{1}{n * \mu_1}$$

$$= \frac{(q-l)_{k+1}}{(N-l)_{k+1}} * \frac{(N)_l}{(q)_l} * \frac{1}{n * \mu_1}$$

$$\leqslant \frac{(q)_{k+1}}{(N)_{k+1}} * \frac{(N)_l}{(q)_l} * \frac{1}{n * \mu_1}$$

(cf. proof of Lemma 3.6)

$$= \left(\frac{q}{N}\right)^{k+1} * \left(\frac{N}{q}\right)^l * \frac{1}{n * \mu_1} * (1 + o(1))$$

$(k, l = o(n^{1/2}))$

$$= \left(\frac{N}{q}\right)^{k+1} * \frac{1}{n * \mu_1} * (1 + o(1))$$

$$\leqslant \left(\frac{2(n-1)}{C}\right)^{k+1} * \frac{1}{(n-1)^k * n * 2^{k-1}} * (1 + o(1)) \qquad (k = o(n^{1/2}))$$

$$\leqslant 4 * \frac{1}{C * n}(1 + o(1)) \qquad (k = \lceil \log_C n \rceil)$$

$$= o(1). \quad ∎$$

The proof of the following lemma would finish the proof of Theorem 3.1.

3.8. Lemma. $EX_{3b}/(EX)^2 = o(1)$, $EX_5/(EX)^2 = o(1)$, $EX_7/(EX)^2 = o(1)$. *The same applies for EX_{4b} and EX_6.*

The cases of Lemma 3.8 require a more detailed analysis of the ways in which the main paths of π' can have common edges with $\pi \cup \bar{\pi}$. We include the lengthy proof in the Appendix.

APPENDIX: PROOF OF LEMMA 3.8

We begin by distinguishing two kinds of chains on $\pi \cup \bar{\pi}$. Note that in 3.4(a) we define a chain of S on V, where V is a path. But $\pi \cup \bar{\pi}$ is not a path. The graph $\pi \cup \bar{\pi}$ has the form of Fig. 12; note $\bar{\pi} = \text{Smp } \pi \cup \text{Ssp } \pi$.

A.1. Definition. Let $S \subseteq E(\pi, \cup \bar{\pi})$ and let $C = e_1 \cdots e_n$, where $n \geqslant 1$, be a sequence of consecutive edges of $\pi \cup \bar{\pi}$:

C is a *chain* of S on $\pi \cup \bar{\pi}$ iff $C \subseteq S \cap E(\pi)$ or $C \subseteq S \cap E(\bar{\pi})$ and if f is a successor of e_n or a predecessor of e_1 on $\pi \cup \bar{\pi}$; then $f \notin S$ and $\bar{f} \notin S$.

Let $\text{Chain}_{\pi \cup \bar{\pi}} S$ be the set of chains of S on $\pi \cup \bar{\pi}$. Let $\text{Chain}_\pi S$ the subset of these chains which are contained in π.

C is a *broken* chain of S on $\pi \cup \bar{\pi}$ iff $C \subseteq S$ and $C \cap E(\pi) \neq \varnothing$ and $C \cap E(\bar{\pi}) \neq \varnothing$ and if f is a successor of e_n or a predecessor of e_1 on $\pi \cup \bar{\pi}$; then $f \notin S$ and $\bar{f} \notin S$.

$\text{Brchain}_{\pi \cup \bar{\pi}} S$ is the set of broken chains of S on $\pi \cup \bar{\pi}$. Broken chains go from π to $\bar{\pi}$ or vice versa, whereas chains either are in π or in $\bar{\pi}$.

A.2. Corollary. (a) *If* (π, π') *satisfies condition* 3.b *with parameters* r, R, *then* $|\text{Chain}_{\pi \cup \bar{\pi}} S| = R$. (*For* S *see Definition* 3.4(b), *condition* 3.b.)

(b) *If* (π, π') *satisfies condition* 5 *with parameters* r, R, *then either* $|\text{Chain}_{\pi \cup \bar{\pi}} S| = R$ *or* $|\text{Chain}_{\pi \cup \bar{\pi}} S| = R - 1$ *and* $|\text{Brchain}_{\pi \cup \bar{\pi}} S| = 1$.

(c) *If* (π, π') *satisfies condition* 7 *with parameters* r, R, t, T *then the following two statements hold*:

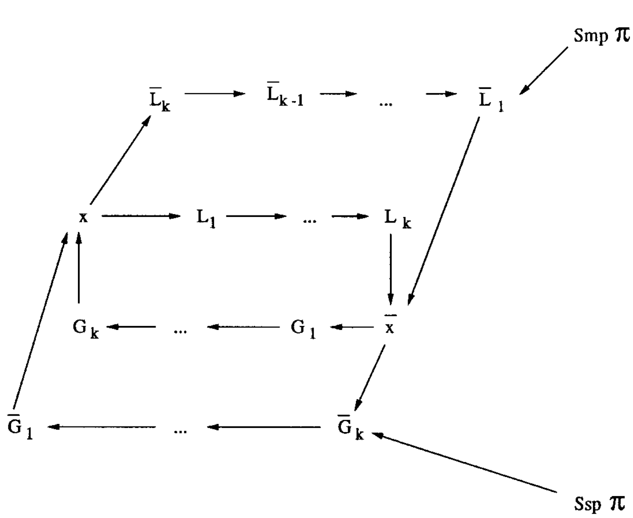

FIGURE 12

First statement: *Either* $|\text{Chain}_{\pi \cup \bar{\pi}} S| = R$ *or* $|\text{Chain}_{\pi \cup \bar{\pi}} S| = R - 1$ *and* $|\text{Brchain}_{\pi \cup \bar{\pi}} S| = 1$.

The second statement is analogous with S', T *instead of* S, R.

Proof. (a) Let $C = e_1 \cdots e_n \in \text{Chain}_{\text{Smp } \pi'} S$. Then C is a sequence of consecutive edges of $\text{Smp } \pi \cup \text{Ssp } \pi$. Let f be a successor or predecessor of C on $\pi \cup \bar{\pi}$. Then $f \notin S$ because otherwise $C \notin \text{Chain}_{\text{Smp } \pi'} S$. Now, assume f is a predecessor of C on $\pi \cup \bar{\pi}$. If $\bar{f} \in S$ we have literals L, G, H such that $e_1 = L \to G$, $\bar{f} = \bar{L} \to H$, and e, $\bar{f} \in E(\text{Smp } \pi')$. This cannot be because a second main path begins at $\bar{x}$ and ends at x without contradictory pairs in between. Therefore $C \in \text{Chain}_{\pi \cup \bar{\pi}} S$. The direction $\text{Chain}_{\pi \cup \bar{\pi}} S \subseteq \text{Chain}_{\text{Smp } \pi'} S$ follows directly.

(b) As in (a) we can show

$$\text{Chain}_{\text{Smp } \pi'} S = \text{Chain}_{\pi \cup \bar{\pi}} S \cup \text{Brchain}_{\pi \cup \bar{\pi}} S.$$

If $|\text{Brchain}_{\pi \cup \bar{\pi}} S| > 1$ then we would have two chains containing segments like $L \to G \to H$ and $A \to \bar{G} \to B$ in $\text{Chain}_{\text{Smp } \pi'} S$. This cannot happen on a second main path, which has contradictory literals only at the beginning and end.

(c) The claim follows as in (b). ∎

A.3. Definition and Corollary. (a) *Let* π *be a simple cycle. For* r, R *with* $1 \leqslant r \leqslant k$, $1 \leqslant R \leqslant r$, *and* $R \leqslant k - r + 2$ *we define*

$$L_\pi(r, R) = |\{S \subseteq E(\pi) \mid |S| = r, \; |\text{Chain}_{\pi \cup \bar{\pi}} S| = R\}|,$$

hence $L_\pi(r, R)$ *is the number of ways in which we can choose* r *edges from* π *such that they form* R *chains on* π;

$$G_\pi(r, R) = |\{S \subseteq E(\pi \cup \bar{\pi}) \mid |S| = r,$$
$$|\text{Brchain}_{\pi \cup \bar{\pi}} S| = 0| \; |\text{Chain}_{\pi \cup \bar{\pi}} S| = R\}|,$$

hence $G_\pi(r, R)$ *is the number of ways to choose* r *edges from* $\pi \cup \bar{\pi}$ *such that they form* R *chains on* $\pi \cup \bar{\pi}$;

$$K_\pi(r, R) = |\{S \subseteq E(\pi \cup \bar{\pi}) \mid |S| = r,$$
$$|\text{Chain}_{\pi \cup \bar{\pi}} S| = R - 1, \; |\text{Brchain}_{\pi \cup \bar{\pi}} S| = 1\}|,$$

hence $K_\pi(r, R)$ *is the number of ways to choose* r *edges from* $\pi \cup \bar{\pi}$ *such that they form* $R - 1$ *chains on* $\pi \cup \bar{\pi}$ *and one broken chain.*

Finally, let

$$M(r, R) = 2^R \cdot \frac{l}{R} \cdot \binom{r}{R} \cdot \binom{l - r}{R}.$$

(b) *We have*

$$L_\pi(r, R) = \frac{l}{R} \cdot \binom{r-1}{R-1} \cdot \binom{l-r-1}{R-1} \qquad (\textit{see also } [16]),$$

$$G_\pi(r, R) \leqslant 2^R \cdot L_\pi(r, R) \leqslant M(r, R),$$

$$K_\pi(r, R) \leqslant M(r, R).$$

Proof. (b) Let

$$\pi = \begin{array}{ccc} x & \longrightarrow & P \\ \uparrow & & \downarrow \\ Q & \longleftarrow & \bar{x} \end{array},$$

where $P = \text{Fmp}\,\pi$ and $Q = \text{Smp}\,\pi$. The first claim: If $r = 1$ then $R = 1$ we have l possibilities and the claim holds.

Now, let $r > 1$. First we determine the number of sets $S \subseteq E(\pi)$ with $|S| = r$ and $|\text{Chain}_\pi S| = R$ such that S contains the edge $x \to L$, where L is the first literal of P.

Each S which neither contains the predecessor of $x \to L$ nor the successor can be obtained in exactly one way by the following choosing process:

Choose the lengths of the chains except $x \to L$; i.e, choose a vector $(v_1, ..., v_{R-1})$ with $v_i > 0$ such that $v_1 + \cdots + v_{R-1} = r - 1$: $\binom{r-2}{R-2}$ possibilities [15, Exercise 11, p. 13] Choose the number of edges between the chains; that is, choose a vector $(w_1, ..., w_r)$ with $w_i > 0$ and $w_1 + \cdots + w_R = l - r$: $\binom{l-r-1}{R-1}$ possibilities. In total we have $\binom{r-2}{R-2} \cdot \binom{l-r-1}{R-1}$ possibilities.

If S contains either the predecessor of $x \to L$ or the successor we have $2 \cdot \binom{r-2}{R-1} \cdot \binom{l-r-1}{R-1}$ possibilities. If S contains both the predecessor and the successor of $x \to L$ we have $\binom{r-2}{R} \cdot \binom{l-r-1}{R-1}$ possibilities. All together this gives $\binom{r}{R}\binom{l-r-1}{R-1}$ possibilities. If S does not contain $x \to L$ we get $\binom{r-1}{R-1}\binom{l-r}{R}$ possibilities in a similar way. And

$$\binom{r}{R} \cdot \binom{l-r-1}{R-1} + \binom{r-1}{R-1}\binom{l-r}{R}$$
$$= \left(\frac{r}{R} + \frac{l-r}{R}\right)\binom{r-1}{R-1}\binom{l-r-1}{R-1} = \frac{l}{R}\binom{r-1}{R-1}\binom{l-r-1}{R-1}.$$

The second claim: Each set S with $|S| = r$ and $|\text{Chain}_{\pi \cup \bar{\pi}} S| = R$ and $|\text{Brchain}_{\pi \cup \bar{\pi}} S| = 0$ can be obtained as follows:

(1) Choose r edges in R chains on π: $L_\pi(r, R)$ possibilities.

(2) For each chosen chain choose its complement or not: 2^R possibilities.

The third claim: Each set S with $|S| = r$, $|\text{Chain}_{\pi \cup \bar{\pi}} S| = R - 1$, and $|\text{Brchain}_{\pi \cup \bar{\pi}} S| = 1$ can be obtained as follows:

(1) Choose $R+1$ chains on π such that one chain begins (ends) with x and one ends (begins) with $\bar{x}$: $\leqslant L_\pi(r, R+1)$ possibilities.

(2) For each of the $R-1$ chains, not contributing to the broken chain decide whether its complement is chosen or not: 2^{R-1} possibilities.

(3) For the two chosen chains on π contributing to the broken chain choose which one to complement: 2 possibilities.

Altogether we have $\leqslant M(r, R)$ possibilities. ∎

A.4. Definition and Lemma. (a) *For r, R with $1 \leqslant r \leqslant k$, $1 \leqslant R \leqslant r$, and $R \leqslant k - r + 2$ let*

$$v_1(r, R) = M(r, R) \cdot \frac{(n-1-r-R)! \cdot k^2}{(k-r-R+2)! \cdot (n-1-k)!} \cdot (k-r)! \cdot 2^{k-r-R+2},$$

$$v_2(r, R) = M(r, R) \cdot \frac{(n-1-k-r-R)! \cdot k^2}{(k-r-R+2)! \cdot (n-1-2k)!} \cdot (k-r)! \cdot 2^{k-r-R+2}.$$

(b) *If $R = 1$, the number of pairs of simple cycles in the complete formula graph satisfying condition* 3.b *with parameters r, R is $\leqslant \mu \cdot n \cdot v_2(r, R)$.*

If $R \geqslant 2$ this number is $\leqslant \mu \cdot n^2 \cdot v_2(r, R)$.

The same applies to condition 4.b.

(c) *If $R = 1$, the number of pairs of simple cycles satisfying condition* 5 *with parameters r, R is $\leqslant \mu \cdot \mu_1 \cdot l \cdot n \cdot v_2(r, R)$.*

If $R \geqslant 2$, this number is $\leqslant \mu \cdot \mu_1 \cdot l \cdot n^2 \cdot v_2(r, R)$.

The same applies to condition 6.

(c) *If $R = 1$ and $T = 1$, the number of pairs of simple cycles satisfying condition* 7 *with parameters r, R and t, T is $\leqslant \mu \cdot l \cdot n \cdot v_1(r, R) \cdot n \cdot v_2(t, T)$.*

If $R = 1$ and $T \geqslant 2$ or if $R \geqslant 2$ and $T = 1$, this number is $\leqslant \mu \cdot l \cdot n \cdot v_1(r, R) \cdot n^2 \cdot v_2(t, T)$; if $R \geqslant 2$ and $T \geqslant 2$ it is $\leqslant \mu \cdot l \cdot n^2 \cdot v_1(r, R) \cdot n^2 \cdot v_2(r, R)$.

Proof. (b) Let $R = 1$. Each pair of simple cycles (π, π') satisfying condition 3.b with parameters r, R can be obtained by the following choosing process:

(1) Choose π: μ possibilities. (This also fixes $\text{Fmp}\,\pi'$.)

(2) Choose the edges from $\pi \cup \bar{\pi}$ which also occur in $\text{Smp}\,\pi'$: $\leqslant M(r, R)$ possibilities by A.3(b) and A.2(a).

(3) Choose the remaining variables for $\text{Smp}\,\pi'$: $\leqslant n \cdot (n-1-k-r-R)! \cdot k^2/(k-r-R+2)! \cdot (n-1-2k)!$ possibilities, which can be seen as follows: Let $x, \bar{x}$ be the contradictory pair of π. Then $x, \bar{x}$ is the contradictory pair

of π'. Let M be the set of nodes touched by the edges chosen in (2). It is easy to see that $|M| = r + R$. We distinguish three cases.

Let $x \notin M$ and $\bar{x} \notin M$. To complete $\operatorname{Smp} \pi'$ we must choose $k-r-R$ variables from $n-1-k-r-R$ variables; hence we have $\binom{n-1-k-r-R}{k-r-R}$ ways to choose.

Let either $x \in M$ or $\bar{x} \in M$. Now we must choose $k-r-R+1$ variables from $n-1-k-r-R+1$ variables. This gives us

$$\binom{n-1-k-r-R+1}{k-r-R+1} \leqslant n \cdot \frac{(n-1-k-r-R)! \cdot k}{(k-r-R+2)! \cdot (n-1-2k)!}$$

ways to choose.

The third case $x \in M$ and $\bar{x} \in M$, cannot occur, because $R=1$ and $\operatorname{Smp} \pi' \neq \operatorname{Smp} \pi$ and $\operatorname{Smp} \pi' \neq \operatorname{Ssp} \pi$. (This is the reason for the fact that the upper bound for $R=1$ is smaller than that for $R \geqslant 2$.)

(4) Order the variables and the chain chosen in (3) and (2): $\leqslant (k-r)!$ possibilities, which can be seen as follows: We distinguish three cases analogously to (3):

Let $x, \bar{x} \notin M$. Then each path $\bar{x} \to P \to x$, where P is made up from the variables and the chain chosen in (3) and (2) uniquely determines a permutation of $k-r-R+R = k-r$ objects. If either $x \in M$ or $\bar{x} \in M$, each path P, as before, uniquely determines a permutation of the $k-r-R+1 = k-r$ variables chosen in (3). The chain chosen in (2) must not be moved because it contains x or $\bar{x}$.

(5) Decide whether the variables chosen in (3) shall be negated or not: $\leqslant 2^{k-r}$ possibilities.

Altogether (1) through (5) give us at most $\mu \cdot n \cdot v_2(r, R)$ ways to choose.

Now let $R \geqslant 2$. Then we have the following choosing process:

(1), (2) as before.

(3) Choose the remaining variables for $\operatorname{Smp} \pi'$: $\leqslant n^2 \cdot (n-1-k-r-R)!\, k^2/(k-r-R+2)! \cdot (n-1-2k)!$ possibilities, which can be seen as follows: As before let $x, \bar{x}$ be the contradictory pair of π and of π'. Let M be the set of nodes touched by the edges chosen in (2). Then $|M| = r + R$. We distinguish three cases.

If $x \notin M$ and $\bar{x} \notin M$ we get the same result as for $R=1$. If either $x \notin M$ or $\bar{x} \notin M$ we get the same result as for $R=1$, too. If $x \in M$ and $\bar{x} \in M$, to complete $\operatorname{Smp} \pi'$ we must choose $k-r-R+2$ variables from $n-1-k-r-R+2$ variables.

Hence we have

$$\binom{n-1-k-r-R+2}{k-r-R+2} \leqslant n^2 \cdot \frac{(n-1-k-r-R)!}{(k-r-R+2)! \cdot (n-1-2k)!}$$

ways to choose.

(4) Order the variables and chains chosen in (3) and (2): $\leqslant (k-r)!$ possibilities, which can be seen by a case distinction analogously to the one made for $R=1$. The only new case which can occur here is: $x \in M$ and $\bar{x} \in M$. Then each path $\bar{x} \to P \to x$, where P is made up from the chosen chains and variables uniquely determines a permutation of $k-r-R+2+R-2 = k-r$ objects. (Not every permutation can occur.) The chain containing x and the chain containing $\bar{x}$ must not be permuted.

(5) The possibility of negating the variables chosen in (3) gives at most $2^{k-r-R+2}$ additional possibilities.

Altogether we have $\leqslant \mu \cdot n^2 \cdot v_2(r, r)$ ways to choose.

(c) Each pair (π, π') satisfying condition 5 with parameters r, R such that for the contradictory pair $y, \bar{y}$ of π' holds $y, \bar{y} \in \operatorname{Lit}(\pi \cup \bar{\pi})$ (note that this does not mean that $y, \bar{y}$ is the contradictory pair of π) is obtained by the first choosing process iff $R=1$ and by the second iff $R \geqslant 2$.

First choosing process: $R=1$.

(1) Choose π: μ possibilities.

(2) Choose $\operatorname{Fmp} \pi'$: $\leqslant l \cdot \mu_1$ possibilities. Note that the factor l allows us to choose the contradictory pair of π' from $\operatorname{Lit}(\pi \cup \bar{\pi})$.

(3) Choose the edges (a single chain or a broken chain) from $\pi \cup \bar{\pi}$ which are also contained in $\operatorname{Smp} \pi'$: $\leqslant M(r, R)$ possibilities.

(4) Choose the remaining variables necessary to complete $\operatorname{Smp} \pi'$: $\leqslant n \cdot (n-1-k-r-R)! \cdot k^2/(k-r-R+2)! \cdot (n-1-2k)!$ possibilities. This can be seen as in (b).

(5) Order the chain and the variables from (3) and (4): $\leqslant (k-r)!$ possibilities, and decide if the variables shall be negated: $\leqslant 2^{k-r-R+1}$ possibilities.

Altogether we have at most $\mu \cdot \mu_1 \cdot l \cdot n \cdot v_2(r, R)$ possibilities to choose.

Second choosing process: $R \geqslant 2$

(1), (2) As in the first choosing process.

(3) Choose the edges from $\pi \cup \bar{\pi}$ or which are also contained in $\operatorname{Smp} \pi'$: $\leqslant M(r, R)$ possibilities.

(4) Choose the remaining variables necessary to complete $\operatorname{Smp} \pi'$: $\leqslant n^2 \cdot (n-1-k-r-R)! \cdot k^2/(k-r-R+2)! \cdot (n-1-2k)$ possibilities. This can be seen as in (b).

(5) Order and decide to negate: $\leqslant (k-r)! \cdot 2^{k-r-R+2}$ possibilities.

Altogether we have $\leqslant \mu \cdot \mu_1 \cdot l \cdot n^2 \cdot v_2(r, R)$ ways to choose.

Each pair (π, π') satisfying condition (5) with parameters r, R such that for the contradictory pair $y, \bar{y}$ of π' holds $y, \bar{y} \notin \operatorname{Lit}(\pi \cup \bar{\pi})$ is obtained by the following choosing process:

(1) Choose π: μ possibilities.

(2) Choose Fmp π': $\leqslant n \cdot \mu_1$ possibilities. Here we have the factor n to allow for the choice of the contradictory pair of π'.

(3) Choose the edges from $\pi \cup \bar{\pi}$ which are also contained in Smp π': $\leqslant M(r, R)$ possibilities.

(4) Choose the remaining variables necessary to complete Smp π': $\binom{n-1-k-r-R}{k-r-R} \leqslant (n-1-k-r-R)! \cdot k^2/(k-r-R+2)! \cdot (n-1-2k)!$ possibilities. As the edges chosen in (2) cannot touch the contradictory pair of π', we must choose $k-r-R$ variables from $n-1-k-r-R$ variables.

(5) Order and decide to negate: $\leqslant (k-r)! \cdot 2^{k-r-R}$ possibilities.

Altogether we have at most $\mu \cdot \mu_1 \cdot n \cdot v_2(r, R)$ ways to choose.

(d) Each pair (π, π') satisfying condition 7 with parameters r, R and t, T such that for the contradictory pair y, $\bar{y}$ of π' holds $y, \bar{y} \in \mathrm{Lit}(\pi \cup \bar{\pi})$ is obtained by:

Choosing process 1 iff $R=1$ and $T=1$,

Choosing process 2 iff $R=1$ and $T \geqslant 2$,

Choosing process 3 iff $R \geqslant 2$ and $T=1$,

Choosing process 4 iff $R \geqslant 2$ and $T \geqslant 2$.

Choosing process 1: $R=1$ and $T=1$

(1) Choose π: μ possibilities.

(2) Choose the contradictory pair of π': l possibilities. Note that we must pick a variable from π.

(3) Choose the edges from $\pi \cup \bar{\pi}$ which are also contained in Fmp π': $\leqslant M(r, R)$ possibilities.

(4) Choose the remaining variables to complete Fmp π': $\leqslant n \cdot (n-1-r-R)! \cdot k^2(k-r-R+2)! \cdot (n-1-k)!$ possibilities, which can be seen as before.

(5) Order and decide to negate: $\leqslant (k-r)! \cdot 2^{k-r-R+1}$ possibilities.

(6) Choose Smp π': $\leqslant M(t, T) \cdot n \cdot ((n-1-k-t-T)! \cdot k^2/(k-t-T+1)! \cdot (n-1-2k)!) \cdot (k-t)! \cdot 2^{k-t+1}$ possibilities.

Altogether we have at most $\mu \cdot l \cdot n \cdot v_1(r, R) \cdot n \cdot v_2(t, T)$ ways to choose.

Choosing process 2: $R=1$ and $T \geqslant 2$:

(1), (2), (3), (4), (5) as in choosing process 1.

(6) Choose Smp π': $\leqslant M(t, T) \cdot n^2 \cdot ((n-1-k-t-T)! \cdot k^2/(k-t-T+2)! \cdot (n-1-2k)!) \cdot (k-t)! \cdot 2^{k-t-T+2}$ possibilities as $T \geqslant 2$.

We have at most $\mu \cdot l \cdot n \cdot M(r, R) \cdot n^2\, M(t, T)$ possibilities.

Choosing process 3: Analogous to choosing process 2.

Choosing process 4: $R \geqslant 2$ and $T \geqslant 2$:

(1), (2), (3) as in choosing process 1.

(4) Choose the remaining variables to complete Fmp π': $\leqslant n^2 \cdot (n-1-r-R)! \cdot k^2/(k-r-R+2)!\,(n-1-2k)!$ possibilities, as $R \geqslant 2$.

(5) Order and decide to negate: $\leqslant (k-r)! \cdot 2^{k-r-R+2}$ possibilities.

(6) Choose Smp π': $\leqslant M(t, T) \cdot n^2 \cdot ((n-1-k-t-T)! \cdot k^2/(k-t-T+2)! \cdot (n-1-2k)!) \cdot (k-t)! \cdot 2^{k-t-T+2}$ possibilities as $T \geqslant 2$.

This gives at most $\mu \cdot l \cdot n^2 \cdot v_1(r, R) \cdot n^2 \cdot v_2(t, T)$ ways to choose.

Each pair (π, π') satisfying condition 7 with parameters r, R and t, T such that for the contradictory pair y, $\bar{y}$ of π' holds $y, \bar{y} \notin \mathrm{Lit}(\pi \cup \bar{\pi})$ is obtained by the following choosing process:

(1) Choose π: μ possibilities.

(2) Choose the contradictory pair of π': $\leqslant n$ possibilities.

(3) Choose the edges from $\pi \cup \bar{\pi}$ which are contained in Fmp π': $\leqslant M(r, R)$ possibilities.

(4) Choose the remaining variables to complete Fmp π': $\binom{n-1-r-R}{k-r-R} \leqslant (n-1-r-R)! \cdot k^2/(k-r-R+2)! \cdot (n-1-k)!$ possibilities. Note that the edges chosen in (3) do not touch the contradictory pair of π'.

(5) Order and decide to negate or not: $\leqslant (k-r)! \cdot 2^{k-r-R}$ possibilities.

(6) Choose Smp π': $\leqslant M(t, T) \cdot ((n-1-k-t-T)! \cdot k^2/(k-t-T+2)! \cdot (n-1-2k)!) \cdot (k-t)! \cdot 2^{k-t-T}$ possibilities to choose.

Altogether we have $\leqslant n \cdot v_1(r, R) \cdot v_2(t, T)$ ways to choose. ∎

The following lemma finishes the proof of Lemma 3.8.

A.5. Lemma. $EX_{3b}/(EX)^2 = o(1)$, $EX_5/(EX)^2 = o(1)$, $EX_7/(EX)^2 = o(1)$: *the same applies to* EX_{4b} *and* EX_6.

Proof. First we decompose the random variables X_i. For r, R with $1 \leqslant r \leqslant k$, $1 \leqslant R \leqslant r$, and $R \leqslant k-r+2$ let

$$X_{3brR} = \sum_{(\pi, \pi')} X_\pi \cdot X_{\pi'},$$

where the sum goes over all pairs (π, π') in the complete formula graph satisfying condition 3.b with parameters r, R and analogously for X_{5rR} and X_{7rRtT}. Then

$$\frac{EX_{3b}}{(EX)^2} = \sum_{r,\, R} \frac{EX_{3brR}}{(EX)^2} = \sum_{\substack{R=1\\ r\geqslant 1}} \frac{EX_{3brR}}{(EX)^2} + \sum_{r\geqslant R\geqslant 2} \frac{EX_{3brR}}{(EX)^2}.$$

The first summand: Let $R=1$. By A.4.(b) we have at most $\mu \cdot n \cdot v_2(r, R)$ pairs (π, π') satisfying condition 3.b with parameters r, R. For each (π, π') we have $\binom{N-l-k-1+r}{q-l-k-1+r}$ graphs in $FG(q)$ containing π and π'. Hence

$$EX_{3brR} \leqslant \mu \cdot n \cdot v_2(r, R) \cdot \binom{N-l-k-1+r}{q-l-k-1+r} \Big/ \binom{N}{q}$$

and

$$\begin{aligned}
\frac{EX_{3brR}}{(EX)^2} &\leqslant \mu \cdot n \cdot v_2(r, R) \\
&\quad \times \binom{N-l-k-1+r}{q-l-k-1+r} \Big/ \binom{N}{q} \times \binom{N}{q}^2 \Big/ \mu^2 \cdot \binom{N-l}{q-l}^2 \\
&= \frac{1}{\mu} \cdot n \cdot v_2(r, R) \cdot \binom{N-l-k-1+r}{q-l-k-1+r} \Big/ \binom{N-l}{q-l} \\
&\quad \times \binom{N}{q} \Big/ \binom{N-l}{q-l} \\
&\leqslant \frac{1}{\mu} \cdot n \cdot v_2(r, R) \cdot \left(\frac{q-l}{N-l}\right)^{k+1-r} \\
&\quad \times \left(\frac{N}{q}\right)^l \cdot (1+o(1)) \qquad (k, l = o(n^{1/2}),\ r \leqslant k) \\
&\leqslant \frac{1}{\mu} \cdot n \cdot v_2(r, R) \cdot \left(\frac{q}{N}\right)^{k+1-r} \cdot \left(\frac{N}{q}\right)^{2k+2} \\
&\quad \cdot (1+o(1)) \qquad \text{(cf. proof of Lemma 3.6)} \\
&= \frac{1}{\mu} \cdot n \cdot v_2(r, R) \cdot \left(\frac{N}{q}\right)^{k+1+r} \cdot (1+o(1)) \\
&\leqslant \frac{1}{n \cdot \mu_1} \cdot n \cdot \left(\frac{2(n-1)}{C}\right)^{k+1} \cdot \frac{1}{\mu_2} \cdot v_2(r, R) \cdot \left(\frac{n}{q}\right)^r \\
&\quad \times (1+o(1))
\end{aligned}$$

as $q = q(n) \geqslant C \cdot n$ and $\mu = n \cdot \mu_1 \cdot \mu_2$.

For the first half of this product we have

$$\begin{aligned}
\frac{1}{\mu_1} \cdot \left(\frac{2(n-1)}{C}\right)^{k+1} &= \frac{1}{(n-1)^k \cdot 2^{k-1}} \cdot \frac{2^{k+1}(n-1)^{k+1}}{C^{k+1}} \cdot (1+o(1)) \\
&\leqslant 4\, \frac{n-1}{C^{k+1}} \cdot (1+o(1)) \leqslant 4 \cdot (1+o(1))
\end{aligned}$$

as $k = \log_C n$. Below we show that

$$\sum_{\substack{R=1\\ r\geqslant 1}} \frac{1}{\mu_2} \cdot v_2(r, R) \cdot \left(\frac{N}{q}\right)^r = o(1).$$

If $r \geqslant R \geqslant 2$ we get in the same manner (note the additional factor n)

$$\frac{EX_{3brR}}{(EX)^2} \leqslant 4 \cdot \frac{1}{\mu_2} \cdot n \cdot v(r, R) \cdot \left(\frac{N}{q}\right)^r \cdot (1+o(1)).$$

Below we show that

$$\sum_{r\geqslant R\geqslant 2} \frac{1}{\mu_2} \cdot n \cdot v_2(r, R) \cdot \left(\frac{N}{q}\right)^r = o(1).$$

Similarly we have

$$\frac{EX_5}{(EX)^2} = \sum_{\substack{R=1\\ r\geqslant 1}} \frac{EX_{5rR}}{(EX)^2} + \sum_{r\geqslant R\geqslant 2} \frac{EX_{5rR}}{(EX)^2}.$$

For each pair (π, π') of simple cycles satisfying condition 5 with parameters r, R we have $\binom{N-2l+r}{q-2l+r}$ graphs of $FG(q)$ containing π and π'. The first summand: Let $R=1$. With A.4.(c) we get

$$\begin{aligned}
\frac{EX_{5rR}}{(EX)^2} &\leqslant \mu \cdot \mu_1 \cdot l \cdot n \cdot v_2(r, R) \\
&\quad \cdot \binom{N-2l+r}{q-2l+r} \Big/ \binom{N}{q} \times \binom{N}{q}^2 \Big/ \mu^2 \cdot \binom{N-l}{q-l}^2 \\
&= \frac{1}{\mu_2} \cdot l \cdot v_2(r, R) \cdot \binom{N-2l+r}{q-2l+r} \Big/ \binom{N-l}{q-l} \\
&\quad \times \binom{N}{q} \Big/ \binom{N-l}{q-l} \\
&\leqslant \frac{1}{\mu_2} \cdot l \cdot v_2(r, R) \cdot \left(\frac{q}{N}\right)^{l-r} \cdot \left(\frac{N}{q}\right)^l \cdot (1+o(1)) \\
&\qquad \text{(cf. proof of Lemma 3.6, } l = o(n^{1/2})) \\
&= \frac{1}{\mu_2} \cdot l \cdot v_2(r, R) \cdot \left(\frac{N}{q}\right)^r \cdot (1+o(1)).
\end{aligned}$$

Let $R \geqslant 2$. Analogously we get (note the factor n)

$$\frac{Ex_{5rR}}{(EX)^2} \leqslant \frac{1}{\mu_2} \cdot l \cdot n \cdot v_2(r, R) \cdot \left(\frac{N}{q}\right)^r \cdot (1+o(1)).$$

For X_7 we have with $A = EX_{7rRtT}/(EX)^2$:

$$\frac{EX_7}{(EX)^2} = \sum_{\substack{r\geqslant R=1\\ r\geqslant T=1}} A + \sum_{\substack{r\geqslant R\geqslant 2\\ t\geqslant T=1}} A + \sum_{\substack{r\geqslant R=1\\ t\geqslant T\geqslant 2}} A + \sum_{\substack{r\geqslant R\geqslant 2\\ t\geqslant T\geqslant 2}} A.$$

For each pair (π, π') satisfying condition 7 with parameters r, R and t, T we have $\binom{N-2l+r+t}{q-2l+r+t}$ graphs in $FG(q)$ containing π and π'.

Let $R=1$ and $T=1$. With A.4(d) we get

$$\begin{aligned}A &\leqslant \mu\cdot l\cdot n\cdot v_1(r,R)\cdot n\cdot v_2(t,T)\\ &\quad\times\binom{N-2l+r+t}{q-2l+r+t}\Big/\binom{N}{q}\times\left(\frac{N}{q}\right)^2\Big/\mu^2\binom{N-l}{q-l}^2\\ &\leqslant\frac{1}{\mu_1}\cdot l\cdot n^{1/2}\cdot v_1(r,R)\cdot\left(\frac{N}{q}\right)^r\\ &\quad\times\frac{1}{\mu_2}\cdot n^{1/2}\cdot v_2(t,T)\cdot\left(\frac{N}{q}\right)^t\cdot(1+o(1)).\end{aligned}$$

If $R=1$ and $T\geqslant 2$ we get

$$\begin{aligned}A &\leqslant\frac{1}{\mu_1}\cdot l\cdot n^{1/2}\cdot v_1(r,R)\cdot\left(\frac{N}{q}\right)^r\\ &\quad\times\frac{1}{\mu_2}\cdot n\cdot n^{1/2}\cdot v_2(t,T)\cdot\left(\frac{N}{q}\right)^t\cdot(1+o(1)),\end{aligned}$$

and analogously for $R\geqslant 2$ and $T=1$.

Finally, for $R\geqslant 2$ and $T\geqslant 2$ we get

$$\begin{aligned}A &\leqslant\frac{1}{\mu_1}\cdot l\cdot n\cdot n^{1/2}\cdot v_1(r,R)\cdot\left(\frac{N}{q}\right)^r\\ &\quad\times\frac{1}{\mu_2}\cdot n\cdot n^{1/2}\cdot v_2(t,T)\cdot\left(\frac{N}{q}\right)^t\cdot(1+o(1)).\end{aligned}$$

Our proof is finished by showing for $i=1, 2$,

$$S_{1,i}=\sum_{r\geqslant R=1}\frac{1}{\mu_1}\cdot n^{1/2}\cdot l\cdot v_i(r,R)\cdot\left(\frac{N}{q}\right)^r=o(1)$$

and

$$S_{2,i}=\sum_{r\geqslant R\geqslant 2}\frac{1}{\mu_i}\cdot l\cdot n\cdot n^{1/2}\cdot v_i(r,R)\cdot\left(\frac{N}{q}\right)^r=o(1).$$

For the following estimates we use some ideas of [16]. First we look at the case $i=2$:

$$\begin{aligned}&\frac{1}{\mu_2}\cdot v_2(r,R)\\ &\quad=2^R\cdot\frac{l}{R}\cdot\binom{r}{R}\cdot\binom{l-r}{R}\\ &\qquad\times\frac{(n-1-k-r)!\cdot k^2}{(k-r-R+2)!\cdot(n-1-2k)!}\cdot(k-r)!\cdot 2^{k-r-R+2}\\ &\qquad\times\left(1\Big/\binom{n-1-k}{k}\cdot k!\cdot 2^{k-1}\right)\\ &\quad=\binom{r}{R}\cdot 8\cdot\frac{1}{2r}\\ &\qquad\times\frac{\begin{array}{c}l\cdot(l-r)!\cdot(n-1-k-r-R)!\\ k^2(k-r)!\,k!(n-1-2k)!\end{array}}{R\cdot R!(l-r-R)!\cdot(k-r-R+2)!}\\ &\qquad\times(1/(n-1-2k)!\,(n-1-k)!\cdot k!)\\ &\quad=\binom{r}{R}\cdot 8\cdot\frac{1}{2^r}\cdot\frac{l\cdot k^2}{R\cdot R!}\cdot\frac{(l-r)_R\,(k-r)_{R-2}}{(n-1-k)_{r+R}}.\end{aligned}$$

Furthermore,

$$\begin{aligned}&\frac{(l-r)_R\,(k-r)_{R-2}}{(n-1-k)_{r+R}}\\ &\quad=\frac{(l-r)_R}{(n-1-k)_r}\cdot\frac{(k-r)(k-r-1)\cdot\;\cdots\;\cdot(k-r-R+3)}{(n-1-k-r)(n-1-k-r-1)\cdot\;\cdots\;\cdot(n-1-k-r-R+3)}\\ &\qquad\cdot\frac{1}{(n-1-k-r-R+2)}\cdot\frac{1}{(n-1-k-r-R+1)}\\ &\quad\leqslant\frac{(l-r)_R}{(n-1-k)_r}\cdot\frac{h^R}{(n-1-k)^r},\end{aligned}$$

where $h=k+2$ as $r+R-1\leqslant k+1$. Note that $k-r-R+2$ and $k-r-R+1$ can become 0 (see the conditions for r, R, k).

The first factor of preceding product can be estimated as

$$\begin{aligned}&\frac{(l-r)_R}{(n-1-k)_r}\\ &\quad=\frac{(l-r)(l-r-1)\cdot\;\cdots\;\cdot(l-r-R+1)}{(n-1-k)(n-1-k-1)\cdot\;\cdots\;\cdot(n-1-k-R+1)}\\ &\qquad\times\frac{1}{(n-1-k-R)\cdot\;\cdots\;\cdot(n-l-k-r+1)}\\ &\quad\leqslant\frac{l^R}{(n-1-k)^R}\cdot\frac{1}{(n-1-2k)^{r-R}}\qquad\text{as}\quad r\leqslant k.\end{aligned}$$

Hence

$$\begin{aligned}\frac{1}{\mu_2}\cdot v_2(r,R) &\leqslant\binom{r}{R}\cdot 8\cdot\frac{1}{2^r}\\ &\quad\times\frac{l\cdot k^2}{R\cdot R!}\cdot\frac{l^R\cdot h^R}{(n-1-k)^{2R}\cdot(n-1-2k)^{r-R}}.\end{aligned}$$

For $i=1$ we get the analogous result with $n-1$ and $n-1-k$, instead of $n-1-k$ and $n-1-2k$ in the denominator.

This implies as $R=1$ in $S_{1,2}$:

$$S_{1,2} \leqslant 8 \cdot \frac{n^{1/2} \cdot l^2 \cdot h^3}{(n-1-k)^2} \cdot \frac{1}{2} \cdot \binom{N}{q} \times \sum_{r \geqslant 1} \binom{r}{1} \cdot \frac{1}{2^{r-1}(n-1-2k)^{r-k}} \binom{N}{q}^{r-1}.$$

As

$$\binom{r}{1} \cdot \frac{1}{2 \cdot (n-1-2k)} \cdot \frac{2(n-1)}{C} = \binom{r}{1} \cdot \frac{1}{C}(1+o(1)) \leqslant \binom{r}{1} \cdot D = \binom{(r-1)+1}{1} \cdot D$$

when n is large and if $D<1$ is a suitable constant (note $n-1-2k \sim n-1$ and $C>1$),

$$\sum_{r \geqslant 1} \binom{(r-1)+1}{1} D^{r-1} = \frac{1}{(1-D)^2}$$

using [11, formula 5.56], and

$$\frac{n^{1/2} \cdot l^2 \cdot h^3}{(n-1-k)^2} \cdot \frac{2(n-1)}{C} = \frac{n^{1/2} \cdot l^2 \cdot h^3}{n} \cdot \frac{2}{C}(1+o(1)) = o(1).$$

Because $l^2 \cdot h^3 = o(n^{1/2})$, we get that $S_{1,2} = o(1)$.

The other sum, $S_{1,1}$, can be treated similarly.

Now, for $S_{2,2}$ we get

$$\begin{aligned} S_{2,2} \leqslant 8 \cdot &\sum_{r \geqslant R \geqslant 2} \binom{r}{R} \cdot \frac{1}{2^r} \\ &\times \frac{l \cdot k^2}{R \cdot R!} \cdot \frac{l^R \cdot h^R}{(n-l-k)^{2R}(n-1-2k)^{r-R}} \cdot \binom{N}{q}^r \cdot l \cdot n \cdot n^{1/2} \\ = 8 \cdot &\sum_{R \geqslant 2} \frac{l \cdot k^2}{R \cdot R!} \cdot \frac{l^R \cdot h^R}{(n-1-k)^{2R}} \\ &\times \frac{1}{2^R} \cdot \binom{N}{q}^R \cdot l \cdot n \cdot n^{1/2} \cdot \sum_{r-R \geqslant 0} \binom{r}{R} \\ &\times \frac{1}{(n-1-2k)^{r-R}} \cdot \frac{1}{2^{r-R}} \cdot \binom{N}{q}^{r-R}. \end{aligned}$$

As before, we have

$$\binom{r}{R} \frac{1}{n-1-2k} \cdot \frac{1}{2} \cdot \frac{N}{q} \leqslant \binom{r}{R} \cdot D = \binom{(r-R)+R}{R} \cdot D$$

where n is large enough and $D<1$ is a suitable constant; hence

$$\sum_{r-R \geqslant 0} \binom{r}{R} \cdot \frac{1}{(n-1-2k)^{r-R}} \cdot \binom{N}{q}^{r-R} \leqslant \frac{1}{(1-D)^{R+r}}.$$

Moreover, for $R \geqslant 2$ we have

$$\begin{aligned} &\frac{l \cdot k^2}{R \cdot R!} \cdot \frac{l^R \cdot h^R}{(n-1-k)^{2R}} \cdot \frac{1}{2^R} \cdot \binom{N}{q}^R \cdot l \cdot n \cdot n^{1/2} \cdot \frac{1}{(1-D)^{R+1}} \\ &\quad \leqslant \frac{1}{(R-2)!} \cdot \frac{l^{R-2} \cdot h^{R-2}}{(n-1-k)^{2(R-2)}} \\ &\quad \times \frac{1}{2^{R-2}} \cdot \binom{N}{q}^{R-2} \cdot \frac{1}{(1-D)^{R-2}} \\ &\quad \times \frac{l^4 \cdot h^4 \cdot n \cdot n^{1/2}}{(n-1-k)^4} \left(\frac{2(n-1)}{C}\right)^2 \cdot \frac{1}{(1-D)^3}. \end{aligned}$$

Then

$$E := \frac{l \cdot h}{(n-1-k)^2} \cdot \frac{1}{2} \cdot \frac{2(n-1)}{C} \cdot \frac{2}{1-D} = o(1)$$

and

$$F := \frac{l^4 \cdot h^4 \cdot n \cdot n^{1/2}}{(n-1-k)^4} \cdot \left(\frac{2(n-1)}{C}\right)^2 \cdot \frac{1}{(1-D)^3} = o(1),$$

because $l^4 \cdot h^4 = o(n^{1/2})$ and $(n-1-k)^4 \sim n^4$.

Finally

$$\begin{aligned} S_{2,2} &\leqslant \sum_{R \geqslant 2} \frac{1}{(R-2)!} \cdot E^{R-2} \cdot F \\ &= e^E \cdot F = (1+o(1))(o(1)) = o(1). \end{aligned}$$

The other sum $S_{2,1}$ can be dealt with similarly. ∎

REFERENCES

1. B. Aspvall, M. F. Plass, and R. E. Tarjan, A linear-time algorithm for testing the truth of certain quantified boolean formulas, *Inform. Process. Lett.* **8** (1979), 121–123.
2. C. A. Brown and P. W. Purdom, An average time analysis of backtracking, *SIAM J. Comput.* **10** (1981), 583–593.
3. M. T. Chao, "Probabilistic Analysis and Performance Measurement for the Satisfiability Problem," Ph.D. dissertation, Case Western Reserve University, 1984.
4. M. T. Chao and J. Franco, Probabilistic analysis of two heuristics for the 3-satisfiability problem, *SIAM J. Comput.* **15** (1986), 1106–1118.
5. V. Chvátal and E. Szemerédi, Many hard examples for resolution, *J. Assoc. Comput. Mach.* **35** (1988), 759–768.
6. O. Dubois Counting the number of solutions for instances of satisfiability, *Theoret. Comput. Sci.* **81** (1991), 49–64.

7. O. Dubois and J. Carlier, Probabilitistic approach to the satisfiability problem, *Theoret. Comput. Sci.* **81** (1991), 65–75.
8. J. Franco, Probabilistic analysis of the pure literal heuristic for solving the satisfiability problem, *Ann. Oper. Res.* **1** (1984), 273–289.
9. J. Franco and M. Paull, Probabilistic analysis of the Davis Putnam Procedure for solving the satisfiability problem, *Discrete Appl. Math.* **5** (1983), 77–87; **17** (1987), 295–299.
10. A. Goldberg, P. W. Purdom, and C. A. Brown, Average time analysis of simplified Davis–Putnam Procedures, *Inform. Process. Lett.* **15** (1982), 72–75; **16** (1983), 213.
11. R. L. Graham, D. E. Knuth, and O. Patashnik, "Concrete Mathematics," Addison–Wesley, Reading, MA, 1988.
12. P. Hansen, B. Jaumard, and M. Minoux, A linear expected-time algorithm for deriving all logical conclusions implied by a set of boolean inequalities, *Math. Programming* **34** (1986), 223–231.
13. I. Palasti, On the threshold distribution function of cycles in a directed random graph, *Studia Math. Hungar.* **6** (1971), 67–73.
14. E. M. Palmer, "Graphical Evolution," New York, 1985.
15. S. Ross, "A First Course in Probability," Macmillan, New York, 1988.
16. E. M. Wright, Large cycles in large labelled graphs, *Math. Proc. Cambridge Phil. Soc.* **78** (1975), 7–17.
17. A. Goerdt, A threshold for unsatisfiability, *in* "Proceedings, MFCS 1992" (J. M. Havel and V. Koubek, Eds.), Lect. Notes in Comput. Sci., Vol. 692, pp. 264–274, Springer-Verlag, New York/Berlin, 1992.
18. V. Chvátal and B. Reed, Mick gets some (the odds are on his side), *in* "Proceedings, 33rd FOCS 1992," pp. 620–627.
19. A. M. Frieze and S. Suen, Analysis of a simple heuristics for random instances of 3-SAT, manuscript, 1993.
20. A. Kamath, R. Motwani, K. Palem, and P. Spirakis, Tail bounds for occupancy and the satisfiability threshold conjecture, *in* "Proceedings, 35th FOCS, 1994," pp. 592–603.
21. A. Broder, A. M. Frieze, and E. Upfal, On the satisfiability and maximum satisfiability of random 3-CNF formulas, *in* "Proceedings, 4th SODA 1993," pp. 322–330.
22. M. Krom, The decision problem for a class of first-order formulas in which all disjunctions are binary, *Math Logik Math.* **3** (1967), 15–20.